DOCUMENT RESUME

ED 286 381 FL 016 931

AUTHOR Nevis, Joel A.
TITLE Finnish Particle Clitics and General Clitic Theory. Working Papers in Linguistics No. 33.
INSTITUTION Ohio State Univ., Columbus. Dept. of Linguistics.
PUB DATE Oct 86
NOTE 171p.; Ph.D. Dissertation, The Ohio State University.
PUB TYPE Dissertations/Theses - Doctoral Dissertations (041)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS Adverbs; *Finnish; *Form Classes (Languages); *Linguistic Theory; *Morphemes; *Morphology (Languages); Phonology; Phrase Structure; Semantics; Structural Analysis (Linguistics); Suprasegmentals; Syntax; Uncommonly Taught Languages
IDENTIFIERS *Clitics; Lappish

ABSTRACT

A doctoral dissertation on general clitic theory and Finnish particle clitics in particular looks at the mixed linguistic status of clitics and proposes that they are not as mysterious as reported, but should be viewed either as bound words (bound lexemes) or phrasal affixes. An initial chapter introduces the Finnish particle clitics and their characteristics and status. A second chapter presents a syntactic account of (sentential) connective and epistemic adverbs and the closely related class of conjunctions, within the framework of generalized phrase structure grammar, to illustrate how Finnish particle clitics are related to independent words that function as sentential adverbs. A third chapter identifies syntactic, semantic, and prosodic properties common to the sentential adverbs and particle clitics, and another chapter reviews the literature on cliticization. Finally, a discussion of the morphosyntactic status of the particle clitics is presented, offering some external evidence for the bound word status of second position clitics. Appended materials include notes on bound words in the Finnish lexicon and a brief comparison of Finnish and Lappish morphology. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED286381

Working Papers in Linguistics

No. 33

FINNISH PARTICLE CLITICS AND GENERAL CLITIC THEORY

Joel A. Nevis

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Dowty

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

October 1986

FL016931

The Ohio State University
DEPARTMENT OF LINGUISTICS

2 BEST COPY AVAILABLE

The Ohio State University

Working Papers in Linguistics # 33

Joel Ashmore Nevis

Finnish Particle Clitics and General Clitic Theory

The Ohio State University

Department of Linguistics

204 Cunz Hall
1841 Millikin Road
Columbus, Ohio 43210-1229
U. S. A.

October 1986

Information about Ohio State University WORKING PAPERS IN LINGUISTICS
(OSU WPL)

Working Papers in Linguistics is an occasional publication of the Department of Linguistics of Ohio State University and usually contains articles written by students and faculty in the department. There are one to three issues per year. Information about available issues appears below. Nos. 1, 5, & 10 are out of print and no longer available.

There are two ways to subscribe to WPL. The first is on a regular basis: the subscriber is automatically sent and billed for each issue as it appears. The second is on an issue-by-issue basis: the subscriber is notified in advance of the contents of each issue, and returns an order blank if that particular issue is desired. To order Working Papers, please send a check made payable to "Ohio State University" to: OSU WPL, Department of Linguistics, Ohio State University, 1841 Millikin Rd., Columbus, OH 43210-1229.

OSU WPL #18, $2.00. 183 pp. (June 1975): Papers by Michael Geis, Sheila Geoghegan, Jeanette Gundel, G. K. Pullum, and Arnold Zwicky.

OSU WPL #19, $2.00. 214 pp. (September 1975), Edited by Robert K. Herbert: Patterns of Language, Culture, and Society: Sub-Saharan Africa contains eighteen papers presented at the Symposium on African Language, Culture, and Society, held at Ohio State University on April 11, 1975.

OSU WPL #20, $2.00. 298 pp. (September 1975), Edited by Robert K. Herbert: Proceedings of the Sixth Conference on African Linguistics contains twenty-seven papers presented at the Sixth Conference on African Linguistics, held at the Ohio State University on April 12-13, 1975.

OSU WPL #21, $2.50. 252 pp. (May 1976), Edited by Arnold Zwicky: Papers on Nonphonology. Papers by Steven Boer and William Lycan, Marian Johnson, Robert Kantor, Patricia Lee, and Jay Pollack.

OSU WPL #22, $3.50. 151 pp. (February 1977), Edited by Olga Garnica: Papers in Psycholinguistics and Sociolinguistics. Papers by Sara Garnes, Olga Garnica, Mary Louise Edwards, Roy Major, and John Perkins.

OSU WPL #23, $3.50. 162 pp. (December 1978): Donegan, Patricia J., On the Natural Phonology of Vowels. (OSU Ph.D. dissertation).

OSU WPL #24, $3.50. 173 pp. (March 1980), Edited by Arnold M. Zwicky. Clitics and Ellipsis. Papers by Robert Jeffers, Nancy Levin (OSU Ph.D. Dissertation), and Arnold Zwicky.

OSU WPL #25, $3.50. 173 pp. (January 1981), Edited by Arnold M. Zwicky: Papers in Phonology. Papers by Donald Churma, Roderick Goman (OSU Ph.D. Dissertation), and Lawrence Schourup.

ISSN 0473-9604

OSU WPL #26, $4.50. 133 pp. (May 1982), Edited by Brian D. Joseph: **Grammatical Relations and Relational Grammar.** Papers by David Dowty, Catherine Jolley, Brian Joseph, John Nerbonne, and Amy Zaharlick.

OSU WPL #27, $4.50. 164 pp. (May 1983), Edited by Gregory T. Stump: **Papers in Historical Linguistics.** Papers by Donald Churma, G. M. Green, Leena Hazelkorn, Gregory Stump, and Rex Wallace.

OSU WPL #28, $5.00. 119 pp. (May 1983). Lawrence Clifford Schourup, Common Discourse Particles in English Conversation. (OSU Ph.D Dissertation)

OSU WPL #29, $5.00. 207 pp. (May 1984), Edited by Arnold Zwicky & Rex Wallace: **Papers on Morphology.** Papers by Belinda Brodie, Donald Churma, Erhard Hinrichs, Brian Joseph, Joel Nevis, Anne Stewart, Rex Wallace, and Arnold Zwicky.

OSU WPL #30, $5.00. 203 pages. (July 1984). John A. Nerbonne, German Temporal Semantics: Three-Dimensional Tense Logic and a GPSG Fragment. (OSU Ph.D. Dissertation).

OSU WPL #31, $6.00. 194 pages. (July 1985), Edited by Michael Geis. **Studies in Generalized Phrase Structure Grammar.** Papers by Belinda Brodie, Annette Bissantz, Erhard Hinrichs, Michael Geis and Arnold Zwicky.

OSU WPL #32, $6.00. 162 pages. (July 1986), **Interfaces.** 14 articles by Arnold Zwicky concerning the interfaces between various components of grammar.

OSU WPL #33, $6.00. 159 pages. (August 1986), Joel A. Nevis, Finnish Particle Clitics and General Clitic Theory. (OSU Ph.D. Dissertation).

The following issues are available through either: The National Technical Information Center, The U.S. Department of Commerce, 5285 Port Royal Rd, Springfield, VA 22151 (PB) or ERIC Document Reproduction Serv. (ED) Ctr. for Applied Linguistics, 1611 N. Kent St., Arlington, VA 22209.

#2. November 1968, 128 pp. (OSU-CISRC-TR-68-3). PB-182 596.
#3. June 1969, 181 pp. (OSU-CISRC-TR-69-4). PB-185 855.
#4. May 1970, 164 pp. (OSU-CISRC-TR-70-26). PB-192 163.
#6. September 1970, 132 pp. (OSU-CISRC-TR-70-12). PB-194 829.
#7. February 1971, 243 pp. (OSU-CISRC-TR-71-7). PB-198 278.
#8. June 1971, 197 pp. (OSU-CISRC-TR-71-7). PB-202 724.
#9. July 1971, 232 pp. (OSU-CISRC-TR-71-8). PB-204 002.
#11. August 1971, 167 pp. ED 062 850.
#12. June 1972, 88 pp. (OSU-CISRC-TR-72-6). PB-210 781.
#13. December 1972, 225 pp. ED 077 268.
Parts of #14. April 1973, 126 pp. ED
#15. April 1973, 221 pp. ED 082 566.
Parts of #16, December 1973, 119 pp. ED

FINNISH PARTICLE CLITICS AND GENERAL CLITIC THEORY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By

Joel Ashmore Nevis, B.A., M.A.

* * * * *

The Ohio State University
1985

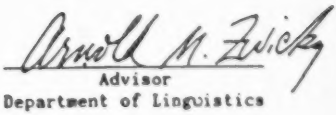
Dissertation Committee:

Arnold M. Zwicky

Ilse Lehiste

Brian Joseph

Approved by



Advisor
Department of Linguistics

ACKNOWLEDGMENTS

I am enormously indebted to my advisor Arnold Zwicky for service above and beyond the call of duty. Arnold was extremely encouraging throughout the writing of this dissertation, and indeed during my entire tenure at OSU. I have benefited from extensive interaction with Arnold as his student, his advisee, and his friend for over six years.

All of my committee members — Arnold Zwicky, Ilse Lehiste, and Brian Joseph -- have been instrumental in shaping my career as a linguist. They have guided both my M.A. thesis and now my Ph.D. dissertation. I thank Ilse Lehiste for promoting Baltic Finnic as a subject of study early in my career and for commenting astutely on my work all along. I am also grateful to Brian Joseph for his critical scrutiny of the present work and for advice during my graduate studies in general.

I thank the American-Scandinavian Foundation for endowing the Finnish Fund Fellowship for the year 1983-84 academic year, as well as the U.S. Educational Foundation in Finland for providing Fulbright Travel Funds. During that year at the University of Helsinki, I collected a great deal of data on the particle clitics. Professor Auli Hakulinen was kind enough to discuss my early ideas and refer me to additional references. She put me onto the parallels between *-hAn* and *sitä* (and the other epistemic adverbs). I have benefitted also from interactions with Professor Fred Karlsson, Professor Raija Bartens, and lecturer Maria Vilkuna, all at the University of Helsinki. Pertti Pyhtilä, also at the University of Helsinki, hereby receives acknowledgement and fond appreciation for his contributions -- sending inaccessible materials, acting as informant, and providing moral support.

Occasionally I have querried other Finns on their language; thanks go especially to the Veikko Kettunen family for their patience in this matter and to Riitta Välimaa-Blum at OSU, who took time to discuss the Finnish data with me (and to proofread part of the manuscript). Her native speaker intuition was of great utility, and her judgments on many sentences helped ward off a number of errors.

Several other members of the OSU Linguistics Department of Linguistics have contributed to the present work. Among them I single out Rich Janda and Jane Smirniotopoulos.

At each stage of researching and writing this dissertation I found my work to be facilitated by the presence and advice of the departmental secretary, Marlene Payha. Marlene helped me wade through administrative paperwork; she instructed me in the use of the Wang word processor; and she saved me a great deal of wasted effort by assisting me through several Wang "dysfunctions". Thanks go also to the person that put ö and ä on the Wang (Epson LQ 1500) -- you simplified my life by two umlauts and a slew of back spacing.

Finally, no dissertation could ever be completed without the emotional (and sometimes financial) support of one's loved ones. My parents, Arnold and Newlin Nevis, provided this plus a whole lot more. They both have always demanded high academic achievement and have acted as role-models in that regard, too. My good friend Jeff Prine put in overtime in the moral support department. I only wish it could have been done without all those long-distance phone bills.

My interest in Finnish clitics arose already in my 1981 master's thesis (Object Case Marking in Finnish), where, besides exploring one aspect of Finnish syntax, I observed in a footnote that the possessive suffixes are really clitics. I returned to this topic later in an article that appeared in OSU WPL 29. Although I am no longer committed to such an abstract analysis as I offer there, the basic arguments to the effect that both the possessive suffixes and the particle clitics have mixed morphosyntactic properties and that the two kinds of clitics are further distinguishable from each other are to be found in that work.

My work in Finnish clisis is intertwined with my research in Estonian clisis. My 1982 CLS contribution established four of the fourteen case endings in Estonian as bound postpositions. As soon as I noticed a discrepancy between my view of the clitic as a bound word and others' views of the notion of clitic, I returned to the Finnish data to argue that the particle clitics are syntactic words despite their phonological dependence (cf. Nevis 1984b). A summary of my ideas can be found in Nevis 1985a, where I discuss several recent articles on Estonian clisis. Chapter IV in the text emulates those ideas.

Material from two other chapters have appeared in print elsewhere. Chapter VI is a revised version of a CLS paper (1985). My ESCOL '85 contribution is adapted mainly from chapter IV.

The ideas and analyses in the present work did not appear suddenly. They have come about through extensive discussion with Arnold Zwicky. Much of this dissertation is shaped by reactions to Zwicky's line of research, which has, over the years, lured me to his theoretical approach. The interface program first appealed to me in a summer seminar of 1981, when I was finishing my M.A. thesis. An early draft of Zwicky and Pullum's book (together with Zwicky's lectures from the summers of 1981 and 1982) opened the topic of clitic theory for me. Zwicky's grammatical model, however, has undergone several changes since then.

My own work has fallen in line with the interface program, with in fact a strengthening effect on the framework. The current version of the Interface Program assumes GPSG as a syntactic approach, which was not in Zwicky and Pullum's earlier draft. I, too, adopt GPSG for a description of Finnish syntax in chapter III, though the results of chapter II are in fact atheoretical. Unfortunately, the major book on GPSG was available too late to be incorporated into this dissertation. Hence certain discrepancies may exist between the version of the theory used here and the revised version.

June 1986 Postscript:

This dissertation is being published as a volume of OSU WPL from the original manuscript with as few alterations as possible. I have not made any substantial revisions, but I have seen to it that typographical errors were corrected. The bibliography was updated to the extent that unpublished manuscripts which have now been published have had the benefit of some additional references. In the original the Derbyshire (1979) reference was inadvertantly left out. It has been added at the end of the document. Two pages have also been removed, namely page v with my vita and page vi with my publications list.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iii

CHAPTER PAGE

I. INTRODUCTION 1

1.0. On Clitics 1
1.1. On Clitic Theory 2
1.2. Organization Of The Dissertation 3
Endnote. 4

II. THE FINNISH PARTICLE CLITICS 5

2.0. Introduction 5
2.1. On The Particle Clitics 5
2.1.1. The Pragmatics/Semantics Of The Particle Clitics ... 5
2.1.2. Other Morphemes That Have Been Analysed As Clitics . 14
2.1.3. The Syntactic Structures Of The Particle Clitics ... 19
2.2. Syntax Vs. The Lexicon 23
2.3. The Clitic Status Of The Particle Clitics 25
Endnotes 36

III. ON FINNISH SENTENTIAL ADVERBS 38

3.0. Introduction 38
3.1. General Remarks On Finnish Sentential Adverbs 38
3.2. Fragment Of Finnish Syntax 40
3.2.1. On Generalized Phrase Structure Grammar 40
3.2.2. Sentential Elements 44
3.2.3. On The Finnish 'Verb Phrase' 45
3.2.4. Phantom Categories 52
3.2.5. The Syntax Of Adverbials 52
3.2.6. Margins in GPSG 61
3.2.7. Topicalization 61
3.3. Summary 64
Endnote 66

IV. TOWARD A SYNTACTIC ACCOUNT OF THE PARTICLE CLITICS 67

4.0. Introduction 67
4.1. Assignment of Phrasal Prominence 67
4.1.1. Stresslessness And Clitichood 68
4.1.2. Prosodic Units 70
4.1.2.1. The Phonological Phrase 72
4.1.2.2. The Phonological Word 72
4.1.2.3. The Finnish Phonological Phrase: Some Problems ... 73
4.1.3. Leaners and Clitics 74
4.1.3.1. Additional Leaners And Clitics In Finnish 76
4.1.4. On Prosodic Theory 77
4.2. Argument For Particle Clitics As Words 77
4.2.1. Liaison 79
4.2.2. On The Notion 'Clitic' 81
4.2.3. Bound Words: A Problem 81
4.3. On The Phrasal Affix 82
4.3.1. Phrasal Affixes: A Problem 84
4.4. Conclusion 88
Endnotes 89

V. ON CLITICS AND CLITICIZATION 91

5.0. Introduction 91
5.1. Clitic Taxonomies 93
5.1.1. Introduction 93
5.1.2. Zwicky 1977 93
5.1.3. Pullum and Zwicky 1983 95
5.1.4. Klavans 1982a 96
5.2. Three Views Of Clisis 99
5.2.1. Clitics As Linguistic Primes 99
5.2.2. Clitics As Affixes 100
5.2.3. Clitics As Words 101
5.3. On The Syntax Of Clitics 103
5.3.1. Heads And Margins 103
5.3.2. Wackernagel's Law 104
5.3.3. The Case Against Kaisse 109
5.4. On The Phonology Of Clitics 109
5.5. Clitic Diachrony 110
5.5.1. Second Position Enclisis 110
5.5.2. On The Agglutination Hypothesis 111
5.6. Summary 113
Endnotes 114

VI. LANGUAGE EXTERNAL EVIDENCE FOR CLITICS AS WORDS: LAPPISH PARTICLE CLITICS 116

6.0. Introduction 116
6.1. Evidence 116
6.2. External Evidence 117
6.3. Evidence From Borrowing 118
6.3.1. The Finnish-Lappish Contact Situation 119
6.3.2. Finnish And Lappish Particle Clitics 120
6.4. Summary 124
Endnotes 125

VII. CONCLUSION 126

7.0. Introduction 126
7.1. Autonomous Components 127
7.1.1. The Principle Of Phonology-Free Syntax 129
7.1.2. The Principle Of Superficial Constraints in Phonology 129
7.1.3. Toward A Restrictive Theory of Cliticization 130
7.2. The Interface Model 132
7.2.1. Zwicky 1982b and 1984c 132
7.2.2. Kaisse 1985 133
7.2.3. Selkirk 1984 133
7.2.4. Revised Organization 134
7.3. Polysynthesis 135
7.4. Concluding Remarks 137
Endnotes 138

APPENDICES

A. Bound Words in the Lexicon 139

B. Comparison of Finnish and Lappish Morphology 146

LIST OF REFERENCES 149

CHAPTER I

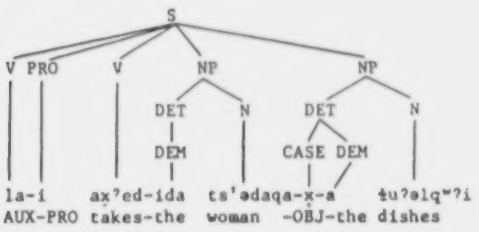
INTRODUCTION

1.0. On Clitics.

The mixed syntactic, morphological, and phonological status of clitics has been the source for varied analyses and raging controversies. One finds an assortment of views on clitics and their status in the grammar, ranging from the suggestion that the set of clitics constituent a linguistic prime, or at least have unusual properties, to the argument that clitics are phrasal affixes or bound words.

The mixed status of clitics can be seen most clearly in the independent syntactic and phonological parameters needed to characterize clitics. Although both parameters frequently run parallel (i.e. syntax and phonology provide the same results), examples like Kwakwala below indicate that the two sometimes diverge. Of special interest here is the case marker in Kwakwala. It forms a syntactic constituent with its following NP but a phonological constituent with the word that precedes it. In (1) the object case marker -x̱ precedes its NP 'the dishes' yet it attaches to the preceding N 'woman' of the subject NP 'the woman'.[1]

(1)



1

In (1) the object case marker attaches to whatever precedes the object NP, not just the subject NP in this VSO language. Example (2) illustrates this point, since the same marker appears on the verb when there is no intervening subject.

(2) k^{w}ix?id-i-x-a q'asa-s-is t'əlwag^{w}ayu
clubbed-he-OBJ-the otter-INSTR-his club
'He clubbed the sea-otter with his club'
(Anderson 1984)

There is no one satisfactory definition of the clitic available in the literature. As Klavans (1982) points out, the correlation between stresslessness and clitichood is imperfect. For this reason I adopt an informal definition: the clitic is a morpheme having a mixed morphosyntactic status, having some word-like characteristics and some affixal characteristics. Zwicky's (1984a) list of diagnostic tests can be used in determining mixed status.

This mixed status of clitics carries over into analyses of cliticization. Many analysts offer cliticization rules that position morphemes syntactically and at the same time attached them phonologically to some host. It is then claimed, contra Zwicky's (1969) Principle of Phonology-Free Syntax, that at least some syntactic rules make reference to phonological information.

1.1. On Clitic Theory.

Several recent proposals concerning clitics offer strong arguments that clisis operations must be distinguished from syntactic operations on the one hand and phonological operations on the other. Such arguments have led to a positing of a separate, autonomous component for cliticization. Syntax is then left without reference to phonological properties. In this work I will explore some of the parameters that have been proposed for the cliticization component. In particular I make some proposals as to how such a component should be restricted in its powers.

In examining the mixed status of clitics, I show that the various properties of clitics are not so mysterious as some analysts have suggested. Instead, I demonstrate that clitics should be viewed either as bound words (bound lexemes) or as phrasal affixes. Thus, sentential clitics generally have a bound word status. Other kinds of clitics may be either bound words or phrasal affixes, depending on the analysis of the facts.

1.2. Organization of the Dissertation.

The information in this dissertation is arranged as follows. Chapter II establishes our object of interest, namely the Finnish particle clitics, and motivates the clitic nature of the particle clitics by demonstrating their mixed syntactic, morphological and phonological status. I show how these morphemes are best viewed as phonologically bound words. Later in the dissertation the particle clitics are related to independent words which function as sentential adverbs (3.0 and 4.2). In order to support this relationship, I present, within the framework of Generalized Phrase Structure Grammar, a syntactic account of (sentential) connective and epistemic adverbs as well as the closely related class of conjunctions (chapter III).

In chapter IV, I connect the sentential adverbs to the particle clitics on the basis of common syntactic, semantic, and prosodic properties. On the one hand, a distinction between merely unstressed words and words that are truly clitic becomes apparent in my examination of prosody in Finnish. The difference between the merely unstressed words (called 'semi-clitics') and true clitics depends on the application of a liaison operation or lack thereof. On the other hand, bound words are formally distinct from phrasal affixes.

In chapter V, I review the literature on cliticization, focussing on the highly relevant proposals by Zwicky, Klavans, and Kaisse. Since the particle clitics are sentential clitics, I address the Kaisse's generalizations over the location of sentential clitics. I return to Zwicky and Pullum's (to appear) Interface Program in the conclusion (chapter VII), where I show how my bound word analysis can be proven to make the framework even more restrictive than, say, Kaisse's (1985) approach. In particular, I argue that the cliticization/readjustment module is made much more uniform and autonomous if it is stripped of all syntactic power.

The issue of the morphosyntactic status of the particle clitics is reviewed in Chapter VI, where I offer some external evidence (namely borrowing) for the bound word status of second position clitics.

The overall goal of this dissertation is a coherent, yet restricted notion of cliticization, one so constrained as to derive almost all properties of cliticization from other aspects of grammar. The theory of cliticization presented in this work is so restrictive as to deny the existence of the clitic *per se*; in lieu of clitics there are bound words (and phrasal) affixes which have direct connections to other phenomena in language.

Endnote.

1. I use the following abbreviations and conventions thoughout this dissertation. For ease of presentation front-back vowel harmonic variants are represented by archiphonemes, with no further theoretical consequence. *A* represents the *a/ä* alternation, *O* the *o/ö* alternation, and *U* the *u/y* alternation. There is no male-female gender distinction in the pronoun *hän*, which means 'he, she'. There are (at least) two infinitives; listed here as the I infinitive and the III infinitive. Note also that my glosses deliberately omit irrelevant grammatical information.

ACC = accusative
ADES = adessive
ADJ = adjective
ADV = adverb, ADVL = adverbial
ALL = allative
COND = conditional
CONJ = conjunction
DER = derivational
EL = elative
EMP = emphasis
FIN = finite
GEN = genitive
H = head
HAN = the *-hAn* morpheme described in section 2.1.1.1.
IFL = informal
ILL = illative
IMP = imperative
INDIC = indicative
INES = inessive
N = noun, NP = noun phrase
NEG = negative
NOM = nominative
NUMB = number
OBJ = object
PASS = passive
PERS = person
PL = plural
POT = potential
PRES = present tense
PRO = pronoun
PX = possessive suffix
Q = interrogative
REFL = reflexive
REL = relative
S = sentence
SG = singular
SITÄ = the morpheme described in section 3.2.5.1.
V = verb

CHAPTER II

THE FINNISH PARTICLE CLITICS

2.0 Introduction.

The particle clitics to be described are *-hAn*, *-pA*, *-kO*, and *-kin/-kAAn*. These have been described previously in terms of their meanings and functions (Karttunen 1975a,b,c; Hakulinen 1976; Östman 1977); their structures have scarcely been touched on (see, however, Penttilä 1963:120 and Kaisse 1982). As a summary I present an approximate semantic/pragmatic characterization of each clitic, as well as a brief description of their distributional facts (section 2.1). In section 2.2, I show why the placement of the particle clitics is a matter for syntax rather than the lexicon. The clitic status of the particle clitics constitutes section 2.3, from which my bound word analysis becomes obvious.

2.1. On The Particle Clitics.

Much of the literature on the particle clitics focuses on their semantics and pragmatics, and as a side issue, on their (morpho)syntactic distribution. I summarize the work on the semantics and pragmatics of the particle clitics in section 2.1.1. In section 2.1.2 I take a look at other morphemes that have been analysed by some as clitics. And in section 2.1.3, I address the problem of the syntactic distribution of the clitics.

2.1.1 The Pragmatics/Semantics of the Particle Clitics.

The meanings of the particle clitics are , first of all, vague. Second, as Hakulinen and Karlsson (1979:327) tell us, their semantics cannot be captured by reference to semantic properties of the morpheme only, but the pragmatics of the sentence must also be taken into consideration, because they all have a textual or interpersonal function. The particle clitics also share a third characteristic: they do not interact with the truth-conditions of the sentence in which they appear. Instead, they give some

additional meaning, such as that of a conventional implication. The importance of these properties will become apparent below in chapter 3 (especially sections 3.0 and 3.1).

2.1.1.1 The Meaning Of *-hAn*.

The clitic *-hAn* has attracted a great deal of attention from linguists in recent years, especially Penttilä, Karttunen, and Hakulinen. The meanings of this clitic are so broad that Penttilä (1957) and the Dictionary of Modern Finnish (NSSK) merely list its various functions. Karttunen (1975a) and Hakulinen (1976) attempt to reduce these different meanings to one basic meaning.

According to Penttilä (1957:120) *-hAn* gives a sentence a flavor that is appealing to the listener (1), mitigating to an expression (2-3), or explanatory of what was said before (4).

(1) Olet-han itsekin samaa mieltä.
are-HAN self-too same opinion
'You are yourself of the same opinion, you know.'

(2) Mitä-hän tuolla tehdään?
what-HAN there does-PASS
'What's being done there, I wonder?'

(3) On-ko-han moisessa perää?
is-Q-HAN such-INES truth
'Is there any truth in something like that, I wonder?'

(4) Hän tuntee minut, on-han hän opettajani
(s)he knows me is-HAN (s)he teacher-my
'She/He knows me, he/she is, after all, my teacher.'

In a similar list, Karttunen (1975a) establishes a plethora of usages -- amelioration (5), contradiction (6), new discovery (7), or a reminder of new truth (8).

(5) Puhu-han asiasta isälle.
talk-HAN matter-EL father-ALL
'Talk to father about it, why don't you.'

(6) Hän ei ole kotona. -- On-pa-han!
(s)he not be home is-EMP-HAN
'He/she is not home. -- She/he is so!

(7) Suomi-han on pieni maa.
Finland-HAN is small land
'Finland is a small country, by golly. (I just found it on the map.)'

(8) Suomi-han on pieni maa.
Finland-HAN is small land
'Finland is a small country, after all. (You don't have to aim all that many rockets at it!)'

But then Karttunen tries to reduce these meanings to one basic, abstract meaning: 'I am in a position to say this to you' (with various extensions which Karttunen lists -- e.g. 'you don't intimidate me, I am reporting something, I have just discovered it, it is common knowledge' and the like). The basic function of *-hAn*, then, according to Karttunen, is the acknowledgement of the speaker's own authority.

Karttunen also considers the role of *-hAn* in subordinated clauses. In embedded sentences, *-hAn* will convey this meaning either from the point of view of the subject of the matrix clause (9), or from the speaker's point of view (10), depending on which verb is used in the matrix clause.

(9) Pekka sanoi että Suomi-han on pieni maa.
Pete said that Finland-HAN is small land
'Pete said that Finland is a small country'

(10) Pekka älysi että Suomi-han on pieni maa.
Pete realized that Finland-HAN is small land
'Pete realized that Finland is actually a small country.'

According to Karttunen (p. 44), in (9) Pete either said the sentence *Pekka sanoi että Suomihan on pieni maa* or he said other things which amount to this statement. In (10), Pete has expressed his own opinion, but the speaker is asserting that Pete is now in agreement with what the speaker already knew.

-hAn is generally a main clause phenomenon. It is permitted in embedded clauses only when the verb of the embedded clause is one of a certain group of semi-factive verbs, as in (10), or when there is a discourse verb, as in (9). Karttunen (1975a) points out that these semi-factive verbs require that the complement clause be true. A discourse verb will refer to the viewpoint of the subject of the main clause, a semi-factive verb the speaker's viewpoint. This is confirmed by negation of the matrix verb (11-12).

(11) Pekka ei sanonut että Suomi-han on pieni maa.
Pete not said that Finland-HAN is small land
'Pete didn't say that Finland is a small country.'

(12) Pekka ei älynnyt että Suomi-han on pieni maa.
Pete not realize that Finland-HAN is small land
'Pete didn't realize that Finland is actually a small country.'

(11) can only be taken as a report that Pete did not happen to utter that exact sentence. (12) means that Pete failed to realize what the speaker knows to be true. Both classes of verbs permit the introduction of new information, whereas other factive verbs such as *tietää* 'to know' generally do not permit this, (13).

(13) *Pekka tietää että Suomi-han on pieni maa.
'Pekka knows that Finland is a small country, after all.'

Hakulinen (1976) likewise begins with a list of the functions of *-hAn* before moving on to one basic, but vague, meaning of the clitic. Her four distinct meanings for *-hAn* include the following:

> 1) appealing to the listener, informing him about an obvious or known fact, or reminding him of what has been said before; 2) coming close to the meaning of an explanatory conjunction; 3) expressing something newly discovered or something that has just happened; 4) making the expression milder a) in statements, implying possibility and doubts, b) to express modesty or indifference, c) in polite questions, d) in polite requests, and e) by means of coming close to the meaning of a concessive conjunction. (Hakulinen 1976:54)

Following Lauri Karttunen, she further reduces these four vague meanings to two basic ones:

> 1) when attached to a verb, *-han* softens up questions, assertions, and commands; 2) when attached to the first constituent of a declarative sentence, *-han* marks something that has just occurred to the speaker, old truths of current relevance, or objection. (Hakulinen 1976:55)

Finally, Hakulinen produces the following single, core meaning for *-hAn*: "to mark a sentence as a reminder of familiar information, and not as a conveyor of new information" (Hakulinen 1976:58). All of the various meanings that have been listed above result from combinations of this core meaning with the basic meaning of the sentence and its possible other implications. As a marker of given or familiar information, *-hAn* cooccurs with the theme of the sentence, but not always attached to it. In sentences containing a theme with no topic, the clitic is located after the theme. But in a sentence containing a topic in addition to the theme, the clitic follows the topic, but precedes the theme.

Hakulinen's approach here is repeated in her other, more recent work (1984a,b), where she emphasizes the textual, cohesive nature of *-hAn* and the other particle clitics.

Välimaa-Blum (1985) takes issue with Hakulinen's (1976) approach. Välimaa-Blum finds it odd that *-hAn* would mark a sentence as containing given information, since new/given

information is already marked by word order in Finnish (i.e. theme precedes rheme). Since the vast majority of sentences contain given information, Hakulinen's approach would actually predict that *-hAn* could be appropriate for the majority of sentences in the language (unless one adopts Hakulinen's 'scale of givenness'). It is not. Instead, it is characteristic only of spoken Finnish -- interpersonal discourse rather than written style (Hakulinen and Karlsson 1979:330; Hakulinen, Karlsson and Vilkuna 1980:121).

Välimaa-Blum takes a somewhat different tack from Hakulinen. She concurs that *-hAn* conveys a conventional implicature, but differs when she emphasizes that *-hAn* indicates new information rather than given. Välimaa-Blum distinguishes between two types of new information: new to the discourse vs. (con)textually new. *-hAn* signals newness of the latter sort. It shows a deviation from an unmarked, neutral textual pattern. She says that the function of *-hAn* is to signal that the "the proposition expressed by the sentence is new with respect to the preceding (non)linguistic context" (1985:1). Thus, Välimaa-Blum is able to take into account the numerous functions of *-hAn*, as well as the restrictions on its occurrence. First, *-hAn* reveals the speaker's attitude toward the proposition as it relates to the rest of the discourse. The absence of *-hAn* indicates a more formal, less interpersonal style. Hence the greater occurrence in spoken dialogue than in written dialogue. Second, it can only cooccur with certain verbs -- e.g. *sanoa, muistaa, huomata, älytä*, and the like -- which are used, by convention, to indirectly convey new information (Karttunen 1975a:15; Välimaa-Blum 1985:6).

In sum, I note some agreement among Hakulinen, Karttunen, and Välimaa-Blum. They all insist that *-hAn* plays a role in text cohesion.

2.1.1.2. The Meaning Of *-kO*.

-kO is the interrogative marker for yes/no questions (11) and for if/whether-type subordinated clauses (15). Its presence is obligatory in both constructions.

(14) On-ko Pertti naimisissa?
is-Q Pertti married
'Is Pertti married?'

(15) En tiedä on-ko Pertti naimisissa.
I-not know is-Q Pertti married
'I do not know whether/if Pertti is married.'

In the neutral yes/no and whether clauses, the finite verb is positioned at the beginning of the sentence and the *-kO* morpheme is attached to it. In topicalized constructions, the topic is placed in clause-initial position and *-kO* is attached to it (16-17):

(16) Pertti-kö on naimisissa?
Pertti-Q is married
'Is it Pertti who is married?'

(17) Naimisissa-ko Pertti on?
married-Q Pertti is
'Is it married that Pertti is?'

2.1.1.3. The Meaning Of -pA.

-pA is usually described as an emphasis marker, as in (18-19). But Hakulinen (1984a:21-22) reports that, in addition, it functions as some sort of interpersonal mitigator, i.e. opening a text (20) (though this example is rather formulaic and not synchronically productive); as an addition of contrastiveness to a focussed element (21); or as a hortative addition to an imperative (22). The following examples have been adapted from Hakulinen 1984a.

(18) On-pa täällä kuuma!
is-EMP here hot
'It sure is hot!'

(19) Se-pä oli yllätys!
it-EMP was surprise
'It certainly was a surprise!'

(20) Oli-pa kerran ...
was-EMP once
'Once upon a time ...'

(21) MINÄ-PÄ siellä kävin.
I-EMP there visited
'It was I who went there.'

(22) Tule-pa tänne.
come-EMP here
'Come here a bit.'

A similar list of functions can be found in Karttunen (1975b), who describes the diversity of the uses of *-pA*: it expresses certainty (23), something just observed (24), intensity in rhetorical questions (as an appeal to common knowledge (25), "you see" as the beginning of stories (26), a 'concessive' (27), or contradiction (28). Examples taken directly from Karttunen (1975b:4):

(23) Kyllä-pä oli hauska.
sure-EMP was fun
'It sure was fun.'

(24) Antti-pa se on.
Andy-EMP it is
'Why it's Andy.'

(25) Kuka-pa ei muistaisi kuinka ...
who-EMP not remember how
'Who wouldn't remember how ...'

(26) Olisin-pa rikas!
would-be-EMP rich
'If only I were rich!'

(27) Oli-pa miten oli.
was-EMP how was
'It was as it was.'

(28) Et saa. Saan-pa.
not get get-EMP
'You may not. Yes, I may.'

The various functions of *-pA* fall out from its basic, but vague, meaning(s) of emphasis and/or exclamation (expressing an extreme or unexpected state of affairs -- Hakulinen 1984a:22) coupled with the context it appears in. Hakulinen (1984b) takes the approach that *-pA* is an intersentential connector with an interpersonal function.

2.1.1.4. The meanings of *-kin/-kAAn*

The meanings of *-kin* and *-kAAn* have been described by a number of analysts, including Östman (1977), Karttunen (1975a-c), Karttunen and Karttunen (1976), Hakulinen and Karlsson (1979), Hakulinen (1984a,b).

2.1.1.4.1 *-kin.*

After *-hAn*, the clitic *-kin* has received the most attention from Finnish linguists. Its basic meaning is 'also, too, even' (Karttunen and Karttunen 1976) as in (29), but it does not have sentential scope under this meaning. In its sentential scope it has a textual function, and attaches to the finite verb (Östman 1977); see (30) below. In this function it indicates something unexpected or, if expected, then as having been learned or being under discussion (Hakulinen and Karlsson 1979:328-30, Hakulinen 1984a:23ff).

(29) Jussi-kin kävi kotona.
Jack -also went home
'Jack, too, went home.'

(30) a. Odotimme sadetta. Tuli-kin pouta.
expected rain came-also fine
'We were expecting rain. But it became nice.'

b. Odotimme sadetta. Sade tuli-kin.
expected rain rain came-also
'We were expecting rain. And the rain did come.'

In (29) *-kin* appears to be a second position clitic (cf. section 2.1.3), but in its local use it can occur anywhere (31, 32).

(31) Jussi kävi-kin kotona.
Jack went-also home
'Jack did too come home.'

(32) Jussi kävi kotona-kin
Jack went home -also
'Jack came HOME, too.'

In its textual function, *-kin* can also indicate emphasis.

(33) [Liisa on todellinen ystävä hädässä.]
'Liisa is a true friend in need.'

Eilen-kin hän teki kaikki kotitehtäväni vaikka en edes pyytänyt
yesterday-kin she did all homework-my though I not even asked
'Just yesterday she did all my homework though I didn't even ask.'

So it seems that this emphatic use indicates that *-kin* may also be a second position clitic.[1] I discuss the significance of the attachment properties of *-kin* in chapter 5.

2.1.1.4.2 -kAAn.

The polar opposite of *-kin* 'also, even' is *-kAAn* 'neither'. The two clitics are in complementary distribution with respect to sentential negation. A change from affirmative to negative in a sentence requires an automatic concurrent change from *-kin* to *-kAAn*. Compare (34) below with (31) above.

(34) Jussi ei käynyt-kään kotona.
Jack not gone-neither home
'Jack didn't go home after all.'

Under certain circumstances the two clitics may stand in contrast with each other, so that there is some sense in which the two morphs constitute separate morphemes (Hakulinen and Karlsson 1979:330).

(35) Ei-kö Swan-37 maksa-kaan 400 000 mk?
not-Q Swan-37 cost-neither 400,000 marks
'The Swan-37 doesn't cost 400,000 Finnmarks, does it?'

(36) Ei-kö Swan-37 maksa-kin 400 000 mk?
not-Q Swan-37 cost-also 400,000 marks
'The Swan-37 does cost 400,000 Finmarks, doesn't it?'

And as Hakulinen and Karlsson (1979:330) point out, the two sentences carry different implicatures. Nonetheless, I treat the

two morphemes as equivalents for the purposes of this dissertation, since the substitution of one for the other is nearly completely automatic.

2.1.1.5. Clitic combinations.

Most of the clitic concatenations are semantically compositional, i.e. their meanings are to be predicted from the meanings of the individual clitics (cf. Karttunen 1975b). A few clitic combinations are not predictable in that way. In particular, the combination of clitics *-pA ... -kin/kAAn* can have a specialized meaning (A. Hakulinen 1984a), as the following example demonstrates. I treat such clitic combinations as discontinuous constituents. This discontinuous clitic compound marks a sentence like that in (37) as a textual mark of "an additional example of something that has been presented before [in the text], and an extreme example at that" (Hakulinen 1984a:22).

(37) [Nämä linnut ovat alkaneet levitä Etelä-Suomeenkin.]
'These birds have begun to spread to South Finland, too.'

On-pa joitakin pariskuntia tavattu Hankoniemen kärjestä kin
is-EMP some pairs met Hanko-peninsula's end
'Some pairs have even been found at the very tip of the Hanko peninsula.'

Similarly, the sequence *-pA-hAn* may have a specialized semantics. Generally *-pA* has an emphatic meaning, but in combination with *-hAn* it is used as a clarifier and appears entirely without emphasis (Hakulinen and Karlsson 1979:330). Since the meaning of the *-pA-hAn* unit is non-compositional, I treat it as a clitic compound.

(38) Ei se mies ollut mikään pelkuri, oli-pa han tavallinen kulkuri.
not that man been any coward, was-EMP-HAN regular vagabond
'That man was not a coward, (he) was a regular vagabond.'

In general it is also possible to predict which clitics may cooccur or may not cooccur on the basis of their basic meanings; see the brief discussions in Karttunen (1975b) and Karlsson (1983). I deal with this matter in section 2.1.3.

Penttilä (1957:556-9) lists other compound combinations -- not clitic compounds, but discontinuous word + clitic compounds. *jopa ... -kin* 'even ... too', *jos ... -kin* 'if ... also', *niin ... kuin ... -kin* 'both ... and also', *niin ... jos ... -kin* 'thus ... if ... too', *niin hyvin ... kuin ... -kin* '... as well as ... too', *vai ... -kO* 'or ... ?'(also *-kO vai*). But to my knowledge these are fully compositional in their semantics and syntax.

In two instances the clitic *-pA* has become lexicalized with another word to form a non-compositional, morphologically frozen unit. First there is *vaikkapa* 'well, say, even', from *vaikka* 'although plus *-pA*. Although *vaikka* is a conjunction, and thus ought to be sentence-initial, and *-pA* likewise is a second position clitic, *vaikkapa* may be used in other positions in the sentence than initial position:

(39) Lähdetään huomenna vaikkapa Pohjois-Karjalaan.
Let's leave tomorrow well North Karelia-ILL
'Let's leave tomorrow, say, for North Karelia. (Any old place will do, but I suggest this one.)'

And second there is *jopa* 'even', from adverb *jo* 'already' plus clitic *-pA*. *Jopa* likewise has a great deal of syntactic freedom (uncharacteristic of true clitic *-pA*) and lacks semantic compositionality.

2.1.2. Other Morphemes That Have Been Analysed As Clitics.

The literature on Finnish particle clitics includes several other morphemes which have been claimed to be clitics. In this section I examine these putative clitics, and in some instances, I argue that the morphemes in question are not, in fact, clitics, but are affixal in nature. In other instances, I come to no conclusions concerning the putative clitics, since I have insufficient data on them.

2.1.2.1. The meaning of *-s*.

The putative clitic *-s* is described as a marker of informality (Karttunen 1975a, 1975c; Hakulinen and Karlsson 1979:227). Again, the contribution of *-s* to the meaning of a sentence can only be pinpointed by taking into account the meaning of the whole sentence. Karttunen (1975c) tells us that

> *-s* is very common in colloquial Finnish, to the point that the omission of an expected *-s* makes a sentence sound brusque and unfriendly rather than merely neutral. ...When strong assertions are made with *-pa*, they usually contain *-s*, too, and the assertion is both direct and personal. Again, how this feels depends on the situation. Generally the omission of *-s* in speech has a negative feeling, but in a situation calling for deference, *-s* could feel rude. Deference is part of pleading, and when an imperative is a plea rather than a command, *-s* is completely excluded, as are *-han* and *-pa*.
> (Karttunen 1975c:235-36)

Karttunen includes *-s* among the particle clitics only because it fits into her clitic (ordering) grid (Karttunen 1975b):

I	II	III
-kin	-ko	-han
-kaan	-pa	-s

The morphemes *-s* and *-hAn* are mutually exclusive. They can both attach to the first constituent of a clause or to another clitic (e.g. *-kO-hAn*, *-kO-s*), but not to each other (**-s-hAn*, **-hAn-s*).

In this dissertation I take the position that *-s* is not a true clitic. I adopt this stand because *-s* exhibits more affixal behavior than the other particle clitics. As I note in section 2.2, the particle clitics are extremely promiscuous in host selection -- they attach to virtually any word class. Affixes, on the other hand, are generally more selective in host attachment (Zwicky and Pullum 1983). For instance, the Finnish allative suffix *-lle* attaches only to nominal word forms -- nouns, adjectives and nominalized verbs, never to finite verbs (**lauloa-lle* 'sing (ALL)'). The morpheme *-s* displays a somewhat restricted host selection, insofar as it cannot attach to just anything. The hosts for *-s* can be listed: the second person imperative verbs, interrogative (and relative) pronouns, and certain particle clitics (*-kO*, *-pA*); see Karttunen 1975c and Penttilä 1957:121. In addition, Penttilä reports that *-s* attaches to *että* and *jotta* in their interjective uses, and to a present passive verb form (in which case the final *n* of the tA-An morphemes is replaced by the *-s*, and the meaning becomes exhortative). The following sentences, from Penttilä 1957:121, exemplify the host requirements:

(41) Second person imperative:

a. Tule-s tänne!
come-IFL here
'Come here!'

b. Olkaa-s nyt hiljaa!
be-PL-IFL now quiet
'Be quiet now!'

(42) Interrogative and relative pronouns:

a. Mitä-s tämä on?
What-IFL this is
'What is this?'

b. Vaimo, jonka-s minulle annoit
wife that-IFL me-to gave-2SG
'The wife that you gave me'

c. Kuinka-s sitten kävi?
how-IFL then went
'How did it go, then?'

(43) *että* and *jotta*:

a. Että-s viitsitte!
that-IFL joke-2PL
"That you are joking!"

b. Jotta-s tiedät!
by-which-IFL know-2SG
"By which you know!"

(44) Particle clitics:

a. Tule-pa-s tänne!
come-EMP-IFL here
'Come here!'

b. Kuka-pa-s sen tietää?
who-EMP-IFL it knows
'Who knows that?'

c. Tämä-kö-s huutamaan?
this-Q-IFL yell-ILL
'And this one [started] to yell?'

In fact, all of the *-s* forms may be seen as lexicalized variants, rather than the result of an *-s* attachment rule. The forms *että*s and *jotta*s can be viewed as variants of *että* and *jotta*. Particle clitics *-pA* and *-kO* have "informal" variants *-pAs* and *-kOs*; the imperative morphemes (2SG) *"-x"* (as described in section 2.3.a) and (2PL) *-kAA* have variants *-s* (where *-s* replaces the *-x* of the 2SG morpheme) and *-kAAs*, respectively; and the interrogative/relative pronouns all have variants ending in *-s*. Lexicalized variants may exist for other words and morphemes as well, e.g. *kyllä*s in (45) as an alternative to regular *kyllä*. And the list may be expanded by the addition of other morphemes (such as the dialectal present passive *-tAAs*).

(45) Kyllä-s sen itsekin hyvin tiedät.
certainly-IFL it self-too well know-2SG
'Certainly you yourself know it well'

In other words, a fairly short list of morphemes that take *-s* can easily be compiled, as in (46). It is not true, however, that under this hypothesis all imperative verbs are to be listed in the lexicon. Instead, *-s* variants are given as part of the imperative morpheme, not as separate entries of word-forms in the lexicon.

(46) 2SG imperative *-s*
2PL imperative *-kAAs*
Interrogative (and relative) pronouns *kukas*, *mikäs*, *jokas*, *kuinkas*, etc. and members of their paradigms
että and *jottas*
Particle clitics *-pAs*, *-kOs*
(and dialectal present passive *-tAAs*)

Apparently no general rule covers the attachment of *-s* -- its "hosts" are a list of idiosyntactically selected morphemes. Since there is no reason that *-s* must be analysed as a clitic, I do not include it with my discussion of the other particle clitics. Instead, I consider it lexicalized.

2.1.2.2. On *-kA*.

Finnish has two morphemes *-kA* which are traditionally considered clitics. I have discussed both in some detail elsewhere (Nevis 1984c). One is the 'lative'[2] *-kA* found attached to interrogative and relative pronouns (e.g. *jonne* -- *jonne-ka* 'whither'). The other is the negative copulative conjunction *-kA*, which attaches to the negative verb, and has the meaning 'and'. It is this latter clitic that interests us now. It might be considered a clitic rather than strictly affixal insofar as it attaches outside the person and number inflections on the negative verb. In (47) *-kä* clearly lies outside the *-n*, *-t*, etc. inflectional morphemes which mark person and number.

(47) en-kä 'and I not'
et-kä 'and you (SG) not'
ei-kä 'and he/she/it not'
emme-kä 'and we not'
ette-kä 'and you (PL) not'
eivät-kä 'and they not'

The syntactic status of *-kA* is implied by its complementary distribution with the regular copulative conjunction *ja* 'and'. The negative verb may cooccur with either *-kA* or *ja*, but not both. In addition, *-kA* does not attach to anything but the negative verb (i.e. not to affirmative verbs, nor to other negative words). In the following, (48) and (49) have the same meaning. (49) is not terribly good (yet not totally ungrammatical either); (50) is clearly ungrammatical.

(48) Lähden ulkomaille en-kä tule takaisin.
go-1SG abroad not-and come back
'I am going abroad and (I am) not coming back.'

(49) Lähden ulkomaille ja en tule takaisin.
'I am going abroad and (I am) not coming back.'

(50) *Lähden ulkomaille ja en-kä tule takaisin.

This morpheme is reported as an independent conjunction in (closely related) Karelian. The difference between the languages, apparently, is the presence of vowel harmony in Finnish (indicating an intimate phonological relationship between the host and *-kA*) and its absence in Karelian (indicating that *ka* is a full word).

In spite of the syntactic function and the morphotactic properties that point to a clitic analysis for *-kA* in Finnish, I exclude this morpheme from my corpus of Finnish sentential clitics. The productivity of the particle clitics (without including *-s*) surely forces us to treat them as syntactic phenomena. By comparison, clitic *-kA* appears in only eleven word-forms -- in the six non-imperative forms of negative *e-* (47) and in its five imperative forms (51).

(51) älä-kä 'and don't (SG)'
älköön-kä 'and don't let him/her/it'
älkäämme-kä 'and let's not, and don't let's'
älkää-kä 'and don't (PL)'
älkööt-kä 'and don't let them'

The small number of word-forms in which *-kA* occurs might easily be accounted for in an inflectional paradigm, rather than being generated via cliticization. Affixal copulative conjunction *-kA* is a relic of a former productive clitic at an earlier stage of the language. Relics of *-kA* also appear in lexicalized word-forms such as *elikkä 'or'*, *ta(h)ikka 'or'*, *vaikka 'although'*, *sekä 'both'*, and *joka* 'every' (Nevis 1985a). In a similar line, Hakulinen and Karlsson (1979:327) say that *-kA* is highly restricted in use, practically lexicalized.

2.1.2.3. Other Clitics.

Penttilä (1957) mentions a few other clitics that occur in Finnish dialects. I list these here for the sake of completeness, but I leave them out of the rest of the dissertation, because I do not have enough information concerning their distribution. So far as I can tell, there are no facts in these dialectal forms that cannot be accounted for in my approach to the other particle clitics.

Some dialects have *-mA* as a clitic. Penttilä (1957:119) cites three instances from various Finnish literary sources (52-54).

(52) härnään-mä häntä huomenna.
tease-I him/her tomorrow
'I'll tease her/him tomorrow.'

(53) ajattelin, että otan-ma-han tuon, sillä muitakaan en saa.
thought-1SG that take-I-HAN that by-which others not get
'I thought that I will take that, since I won't get any others'

(54) kiertelen-mä-hän vielä muutamia vuosia.
circle-I-HAN still several years
'I will circle for still several years.'

Penttilä also cites *luulenmä* 'I think' < *luulen m(in)ä* lit. 'think I' (and its variant *luulenmä*), which is to be found in Standard Finnish -- presumably as a lexicalized entry. There are, to be sure, other such lexicalized words in Finnish, e.g. *näenmä* 'I see' < *näen m(in)ä* lit. 'see I'. In Penttilä's examples, the clitic *-mA* appears in sentential second position, as do other particle clitics *-hAn, -pA, -kO*, but in these examples it also appears attached to the finite verb, as does clitic *-kin/-kAAn*. Its meaning is not, however, that of a 'particle', but pronominal in nature. I do not, therefore, consider *-mA* a sentential clitic. Instead, it would be easy to view *-mA* as a cliticized version of the pronoun *mä* (a shortened variant of *minä*) when it satisfies certain conditions on syntactic positioning (either in 2P or after the finite verb, or both). Even if it were to turn out that *-mA* is a sentential clitic, it would not be a problem for my approach, since it would then be just another sentential second position clitic.

Penttilä (1957:119) mentions another clitic, *-stA*, which (in contrast to *-mA*) apparently is a sentential clitic. It is the clitic counterpart to the adverb *sitä* discussed in section 3.2.5. Again, *-stA* is not included in my corpus of particle clitics, due to my ignorance of its distribution facts.

2.1.3. The Syntactic Structures of the Particle Clitics.

The morphemes *-hAn, -pA, -kO* are enclitic to the first constituent of the sentence. This well-known principle for locating sentential clitics is called **Wackernagel's Law**, after Jacob Wackernagel's (1892) classic exposition of (ancient) Indo European sentential clitic (and 'quasi-clitic') particles, though I prefer to use the term 'second position' to refer to this slot. Second position can mean different things in different languages -- for Ancient Greek second position means after the first word (Kaisse 1985:80), for Finnish it means after the first constituent, and for Luiseño it can mean either (Kaisse 1985:85ff, Pullum 1981: Steele et al. 1981).

Although it is true that *-hAn, -pA, -kO* occur after the initially positioned constituent of a clause, attached to the last word of that constituent, as in nearly all of the examples above, under certain circumstances they attach to some other word of this preceding constituent.

First, when some material intervenes between the head noun of an NP and the clitic, the clitic will attach not to the last word of the preceding constituent, but to the head noun. This situation can be found when a relative clause follows its head, rather than preceding it, as in (55).

(55) vanha mies-hän, joka saapui eilen ...
old man -HAN who arrived yesterday
'The old man, who arrived yesterday ...'

In (55), clitic *-hAn* is attached not to NP-final *eilen*, but to head noun *mies*.

The second situation in which second position clitic attachment is not to the last word of the initially positioned constituent involves a WH-pronoun[3] in an NP, as (56).

(56) Missä-hän maassa me olemme nyt?
what-HAN land we are now
'In what land are we now, I wonder?'

*? Missä maassa-han me olemme nyt?

It is not clear to me whether or not this second exception is the result of topicalization. Topicalization is responsible for the difference between (57) and (58) below, where it looks as if there is optional attachment in an NP between the head noun *auton* and the adjective *uuden*.

(57) Uuden auton-han hän osti
new car-HAN s/he bought
'It was a new CAR that he/she bought.'

(58) Uuden-han auton hän osti
new -HAN car s/he bought
'It was a NEW car that he/she bought.'

I show in chapter 3.2.7 that topicalization is a syntactic rule that makes the topic an immediate daughter of the S node, so that a topic like *uuden* in (58) above will count as the first constituent in the clause. I will henceforth assume that the WH-word is an instance of topicalization.

In examples of topicalization of a subconstituent out of an NP, such as (58), the rest of the NP is generally not separated from its topicalized part in the clause; this is just a tendency in the language, since there are examples in which the NP can be broken up. Compare Karttunen's (1975a:41) examples, repeated here as (59a-b), in which the adverb *kovin* freely leaves its NP for fronting. Admittedly these examples are not very good, but not completely ungrammatical either..

(59) a. Sinä olet vielä kovin nuori.
you are still quite young
'You are still quite young.'

b. Kovin sinä olet vielä nuori.
quite you are still young
'You are still AWFULLY young.'

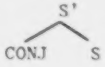
Attachment to the head and to a WH-word is due to a special syntactic operation. I will present an account of these morpheme-peculiar facts in chapter 4.2. Note here that the default instance in cliticization is positioning after a constituent, rather than within a constituent.

In sum, then, the syntactic side of clitic positioning places *-hAn*, *-pA*, and *-kO* in clausal second position (i.e. after the first constituent of the clause, or in the special instances, within an initially located NP). The phonological side of cliticization will merely require the attachment of these clitics to the word that immediately precedes it.

The above facts hold for second position clitics *-hAn*, *-pA* and *-kO*, but not for *-kin/-kAAn*. This last clitic attaches to the finite verb in its sentential use -- see chapter 5.3.

There are also some facts which indicate that the attachment of the particle clitics can be blocked. The particle clitics will not attach to conjunctions, because conjunctions are not actually part of the clause in which the clitics occur. The second positioned particle clitics attach to an initially positioned sister in its S domain. For conjunctions I assume a different structure:

(60)



This assumption accounts for the unacceptability of (61), where the clitic *-hAn* is attached to the conjunction *mutta*, as well as for the acceptability of (62), where *-hAn* attaches to the first constituent *Jussi* in its S domain.

(61) *Mutta-han Jussi ei ole sairas.
but -HAN Jack not be sick

(62) Mutta Jussi-han ei ole sairas.
but Jack -HAN not be sick
'But Jack is not sick after all.'

These attachment facts can be used to determine whether a word is a true conjunction or just a semantically related connective adverb. In (63), for instance, the word *miksi* is not a conjunction, since it behaves as a daughter of S rather than a sister of S for the purposes of clitic attachment.

(63) Miksi-hän Jussi ei ole sairas?
Why -HAN Jack not be sick
'Why isn't Jack sick, after all?'

A second factor that plays a role in blocking is semantic compatibility. An example of this is conjunction-like *jos*, which accepts attachment of *-pA*, but not *-hAn* or *-kO*:

(64) Mä en tiedä jos sinne voisi jo mennä.
I not know if there could already go
'I do not know if one could go there already.'
(a) ...*jos-han ...
(b) ... jos-pa ...
(c) ...*jos-ko ...

These clitics are comparatively rare in subordinated clauses, so it is not an easy task to find good examples of subordinators cooccurring with particle clitics. In (64b), *jos* plus clitic *-pA* is acceptable; in (64a) and (64c) the *jos* plus clitic combination is not. In (64a) this unacceptablity arises because *-hAn* requires that the subordinated clause be true (cf. 2.1.1.1) -- *(en tiedä) jos* '(I do not know) if' indicates that the truth of the clause is in doubt. In (64c) both *jos* and *-kO* are subordinators, and cannot therefore occur in the same clause (though *-kO* is pleonastic in connection with words elsewhere).

Kaisse (1985)[4] notes that second position cliticization interacts with another cliticization rule, namely negative cliticization onto certain complementizers (65).

(65) a. Mutta Jussi ei ole sairas.
but Jack not be sick
'But Jack is not sick.'

b. Muttei Jussi ole sairas.
'But Jack is not sick.'

Other examples include the complementizers in (66)[5].

(66) vaikkei < vaikka ei 'although not'
ettei < että ei 'that not'
jollei < jos ei 'if not'
[*ellei* is a variant of *jollei*.]

When the negative verb (here *ei*) serves as the host for a second position clitic (e.g. *-hAn*), negative cliticization is not permitted.

(67) a. Mutta ei-hän Jussi ole sairas.

b. *Mutt'ei-hän Jussi ole sairas.

[(67b) is unacceptable with main stress on the *mut-* syllable, but completely acceptable with main stress on the *ei* syllable. I comment on this below.]

Kaisse (1982:11) points out that connective *miksi* does permit both cliticizations:

(68) a. Miksi Jussi ei ole sairas?
why Jack not be sick
'Why isn't Jack sick?'

b. Miksei-hän Jussi ole sairas?'

Miksi is not a true complementizer, but is instead the first word of the clause in which the clitic occurs, as in (69). *Miksei* is likewise different from the examples above, (67b) for instance, because it behaves like a single morpheme rather than a contraction of a complementizer and a following negative verb. This fact is clear from stress assignment, according to which we expect main stress on the initial syllable of a word. *Mutteihän* is acceptable only if stress falls on the negative verb *ei* (i.e. *mutteíhän*), but not if stress falls on the initial syllable (**mútteihän*); hence the cliticization of the separate words *mutta* and *ei* is revealed. In comparison, *mikseihän* has stress only on the initial syllable (i.e. *míkseihän*), which points to a monomorphemic analysis for the word *miksei*.[6] Thus *mikseihän* is not due to a rule of negative cliticization.

(69) Miksi-hän Jussi ei ole sairas?
'Why isn't Jack sick, I wonder?'

The facts and analysis I offer above are just suggestive of work that needs to be done on the interaction of complementizers (i.e. conjunctions and subordinators) and/or connective adverbs with particle clitics and negative cliticization. I believe that, if semantic compatibility is present, the location of the particle clitics in a clause can be used to determine the conjunction/connective status of a word. In other words, if a particle clitic like *-hÄn* attaches to the word in question, that word is a connective; if the clitic attaches to the first constituent after the word in question (e.g. (67a) above), then that word is a conjunction.

2.2. Syntax vs. the Lexicon.

The relevance of syntax to a description of these particle clitics needs to be established. I will argue in section 2.3 that the particle clitics of Finnish are lexemes rather than inflectional or derivational affixes. Here, however, I demonstrate that the particle clitics are not to be listed in the lexicon alongside their hosts. They are bound morphemes, but their positioning and attachment is not due to morphology as in the case of regular affixation or compounding. The host+clitic construct is, instead, a result of syntax and cliticization.

Usually when analysts encounter a bound morpheme, they make the assumption that it is an affix unless otherwise demonstrated (Zwicky 1984a:152). In the case of the particle clitics, we can indeed demonstrate on three grounds that the morphemes in question are not affixes. First, there is the problem of the near-absolute productivity in host+clitic combination -- the clitics are promiscuous in host selection. Second, one must make a statement about the linear positioning of the morpheme in the sentence, something morphology never does. And third, one must be able to handle optional locations within the phrase.

The particle clitics exhibit extreme productivity in host selection; they combine with almost anything (though with certain semantic restrictions). In this regard, clitic+host combinations are much more like word+word combinations than like affix+stem combinations (Zwicky 1984a:152). Inflectional affixes have word-class restrictions on the stems to which they attach, and often display gaps in combining with certain stems (Zwicky and Pullum 1983). But in the case of the particle clitics there are no arbitrary gaps -- only gaps predictable from the semantic properties of the lexemes in question. If the particle clitics were to be listed in the lexicon with their hosts, every word in the language would be listed several times, first alone, then with one or more of the clitics.

Thus alongside *Helsinki*, to take one example, there would be *Helsinki-kö, Helsinki-hän, Helsinki-pä, Helsinki-kö-hän, Helsinki-pä-hän, Helsinki-kö-s*, etc. Then, in the genitive: *Helsingin, Helsingin-kö, Helsingin-hän, Helsingin-pä*, and so forth for all thirteen productive cases and two numbers. Karlsson (1983:356, 1984a:3, 1984b:4) has pointed out that if particle clitics were included in the paradigm for nouns, the number of word-forms in each paradigm would expand to around 2000. For verbs the number would be approximately 18,000 (Karlsson 1983:356-7, 1984a:3-4). The promiscuity of clitics with respect to host selection is apparent here.

Following Zwicky (1985a) we can distinguish between inflection and cliticization on the one hand and derivation on the other:

> Rules of [inflection and cliticization] are purely realizational, while the rules of word formation involve, in addition to morphophonological operations, principles of semantic interpretation and two types of morphosyntactic conditions: input conditions, on the base(s) to which a rule applies; and output conditions, specifying the category and morphosyntactic features of the word formed by the rule. (Zwicky 1985a:fn 3)

The vast numbers of word-forms in the extended nominal and verbal paradigms are indicative of the clitics' promiscuity in host selection. And there are (for these clitics) no morphosyntactic

input and output conditions, nor are there any special semantic interpretation principles needed (as in word formation rules).

One must still make a statement about linear ordering within the sentence. The exact positioning of the clitics has to make reference to the sentence as a whole, because the clitics *-hAn*, *-pA*, and *-kO* are placed in the second position of the sentence, attached to the first constituent.[7] Such reference to syntax is not common for an affix. Further, in the GPSG framework I adopt in chapter 3, syntactic rules may not have recourse to information of the morphological structure of a word.

The third problem is the optional location of the clitic within the phrase. There is typically strict ordering of morphemes within the word, but looser ordering within the phrase (Zwicky 1977:7). Karttunen (1975a:42) offers the following example of optional location within the phrase:

(70) a. Nämä monet hyvin kauniit omenat han
These many very pretty apples-HAN

b. Nämä-hän monet hyvin kauniit omenat
These-HAN many very pretty apples

In (70) the clitic *-hAn* may attach either to the last word in the phrase (here, the head noun) or to the initial determiner of an NP.

I have presented three reasons to exclude clitic attachment from the lexicon: host-clitic combinations are too productive, too decomposable, and too numerous to list in the lexicon. In addition, clitic groups function like full words in their syntax (though not, of course, in their phonology). It is therefore not likely that productive, decomposable complex phrases are to be listed in the lexicon alongside other lexical entries.

Now that I have demonstrated that the placement of the second position clitics is a matter for syntax, I have to determine the syntactic status of these morphemes. This is the subject of the next subsection.

2.3. The Clitic Status of the Particle Clitics.

In syntax, there is (at least) a distinction between the lexeme and the affix. Affixes, I assume hereafter, are assembled in the syntax as feature complexes, but not as constituents in phrase structures, whereas lexemes do appear as constituents. This binary distinction correlates well with many observations about the behavior of lexemes and affixes. Lexemes may act as governors, affixes may not. Features sometimes are determined via agreement; words typically are not.

Following Matthews (1972,1974) and Zwicky (1985a), I assume that morphosyntactic properties are distributed in the syntax of grammar as **features**, and that the morphosyntactic features are realized (via **rules of exponence**) as morpho(phono)logical processes in the morphological component. (Bundles of) morphosyntatic features may correspond to the morphs/processes that realize them in a one-to-one fashion, but not necessarily so. According to Matthews (1972:ch. 6) there can be fused and cumulative exponence, as well as extended and overlapping exponence. There can be, then, many-to-one correspondence, one-to-many correspondence, one-to-none, none-to-one, etc. in the match-up of morphosyntactic features and morphophonological operations.

Now, in order to determine the status of a morpheme, we turn to the tests given by Zwicky and Pullum (1983) and by Zwicky (1984a). Zwicky and Pullum (1983) offer criteria to distinguish the behavior of an affix from that of a clitic. Zwicky (1984a) focuses on distinguishing a clitic from a word. Both works utilize tests rather than definitions -- that is, characteristic symptoms of words and affixes as opposed to invariant properties. Clitics, it will be seen, lie partway between words and affixes, exhibiting some word-like and some affix-like symptoms.

In this section I compare the particle clitics to full words on the one hand and to affixes on the other, demonstrating the mixed status of the morphemes in question. Section 2.4 contains two interpretations of the results: the treatment of clitics as a theoretical construct (separate from the grammatical word, or lexeme, and the affix), or as a non-basic category (bound word vs. phrasal affix).

Clitics exhibit some word-like properties and some affixal properties. One can contrast clitics and words by using the tests listed in Zwicky 1984a. I apply the tests to the particle clitics and, by way of comparison, to the inflected word *korjasi* 'repaired, fixed'. When reference to an inflectional affix is needed, I use the illative (*-lle*). Words may combine into phonological and prosodic units called **phonological phrases**, whereas clitics and affixes combine with lexemes into **phonological words**. In the following I contrast the phonological word with the phonological phrase (a-c).

a. Internal/external sandhi

Internal sandhi rules apply only within phonological words; external sandhi rules apply within phonological phrases. A morpheme affected by or conditioning an internal sandhi rule is affixal in nature; one that is affected by or conditioning a rule of external sandhi is an independent word.

Finnish has several relevant sandhi rules -- word-initial gemination and word-final t-assimilation are good examples of

external sandhi phonological rules, and are indicative of word boundaries rather than clitic or affix boundaries. Stem formation, on the other hand, is a clear indication of word-internal sandhi. Morphemes that are subject to this rule belong to the phonological word; morphemes that are not subject to it belong to separate words.

Word-initial gemination takes place between adjacent full words, but never between a stem and its affix. It is a morphophonemic (or 'morphophonetic' since it affects aphonematic glottal contraction ['] and glottal stop [?] as well) rule that will geminate the initial consonant of a word following certain morphemes and morpheme classes. It is often indicated by a raised /^{X}/ as represented by vene^{X} 'boat', mennä^{X} 'to go', tule^{X} 'come!', tietysti^{X} 'of course', talonsa^{X} 'his/her/their house'. For example, in the phrase vene tulee 'the boat comes', the final /^{X}/ of vene triggers gemination of the /t/ in tulee, and we get [venet:ule:]. On the subject of word-initial gemination, see Karlsson 1983:348ff and references therein.

In this regard, the particle clitics behave like independent words. We find [venek:in, men:äk:ö, tulep:a, tietystik:in, talonsah:an] rather than *[venekin, men:äkö, tulepa, tietystikin, talonsahan].

Word-final t-assimilation, an optional rule, likewise applies across word boundaries (e.g. *miehet kuolivat* [miehek:uolivat] 'the men died', *sanonut sinulle* [sanonus:inul:e] 'said to you') but not within words (e.g. *pitkä* *[pik:ä] 'long', *jo-t-ka* *[jok:a] 'which (PLURAL)'). Again, the particle clitics pattern with the independent words by permitting t-assimilation: *en tiennytkään* [en tien:yk:ä:n]. This ability to undergo external sandhi rules constitutes a strong argument that the particle clitics are full words.

b. Prosodic phonology.

Zwicky (1984a) states:

> Rules of sandhi affect segmental features. But rules of prosodic phonology--rules assigning accent, tone, or length--can also be sensitive to the distinction between phonological words and phonological phrases, in that the domain within which a prosodic feature is distributed can be either the phonological word or the phonological phrase (or some other prosodic unit, like the syllable). Consequently if an element counts as belonging to to a phonological word for the purposes of accent, tone, or length assignment, then it ought to be a clitic rather than a word on its own. And if an element counts as belonging to a phonological phrase for these purposes, it ought to be an independent word rather than a clitic. (Zwicky 1984a:151)

The accentual pattern of the Finnish phonological word is one of alternating stress, starting with initial primary stress; secondary stress falls then on every other syllable except the last, e.g. *käytännöllisèssä* 'useful INESSIVE', *kírjoitèttavìssa* 'being written'. (Certain exceptional patterns are ignored here for the sake of simplicity -- see Karlsson 1983:151). In this instance the particle clitics behave as proper subparts of words:

(71) a. Péruna (*pérunà) kásvaa. 'The potato grows.'
b. Pérunàhan kásvaa. 'The potato grows, you know.'
c. Lúntakòhan? 'Some snow?'

In (71a) **pérunà* is wrong because secondary stress is falling on a word-final final syllable. When a particle clitic is attached to the word, then the *-na* syllable is no longer word-final, so secondary stress is acceptable in (71b). Multiple instances of particle clitics behave as if they were proper subparts of words -- the third syllable, nonfinal *ko* in (71b) has secondary stress just like a veritable affix.

From this test we should conclude that the particle clitics are affixes.

c. Word/phrase domain in segmental phonology.

By this Zwicky intends phonological rules that affect segmental features, yet at the same time, are prosodic in nature because their domains of applicability are prosodic domains. Finnish has one such rule -- vowel harmony -- whereby vowels in a phonological word must agree in tongue advancement (with two 'neutral' vowels /i e/), operating in a left-to-right fashion. Vowel harmony does not apply across words, only within words:

(72) a. tule-ssa -- *tule-ssä 'in the fire'
b. häne-ssä -- *häne-ssa 'in him/her'

but:

c. tulee hän (-- *han) 'comes he/she'

The particle clitics act as proper subparts of words for the purposes of vowel harmony:

(73) tulee-han -- *tulee-hän 'comes'
tulee-ko -- *tulee-kö 'comes Q'
tule-pa -- *tule-pä 'do come!'
ei tullutkaan -- *ei tullutkään 'didn't even come'

Again, the particle clitics behave as if they were affixes.

d. Binding.

Tests d-g are morphological tests which rely on observations about the different morphological behavior of affixes and words.

Zwicky states:

> We expect bound elements to be affixes, free elements to constitute independent words. (1984a:152)

As bound elements, the particle clitics are more affixal and less word-like than independent words. This can used as supportive evidence only, since not all bound elements are affixes (cf. the *cranberry* morphs cited in Aronoff 1976:10ff).

e. Closure.

According to Zwicky it is typical that certain inflectional affixes 'close off' words to further affixation. An element that closes off combinations to affixation is an affix or a clitic rather than a full word. In the Finnish phonological word, case endings close off words to further (derivational and inflectional) affixation, but not to clitics (i.e. possessive suffixes or particle clitics):

ROOT - DERIVATION* - NUMBER - CASE - POSSESSIVE SUFFIX - CLITIC*

(The asterisks here indicate possible multiple occurrences of a particular ending class.) In this scheme, CASE closes off inflectional (and derivational) affixation, POSSESSIVE SUFFIX closes off the rest of inflectional morphology, and CLITIC closes off the rest of morphological combination.[8] No bound (inflectional or derivational) morpheme may be added after CLITIC. Thus CLITIC can be seen to cause absolute closure (while POSSESSIVE SUFFIX triggers morphological closure and CASE causes affixal closure .

This test is not conclusive, but under it the particle clitics are compatible with a word analysis.

f. Construction.

In Finnish, inflectional affixes combine with stems, whereas words combine with other words or phrases. An element whose distribution is stated in terms of its ability to combine with stems is an affix. An element that combines with (potentially) multi-word phrases is an independent word. Because the particle clitics do not combine with stems, and because they do attach to fully formed words and phrases, the particle clitics are clearly word-like on this criterion, as can be seen from the following allomorphy facts.

Stem allomorphy is conditioned by affixes, never by full words. Of the stem types below, only those in the first column combine with words, while those in column two combine with affixes (cf. the allative in column three):

WORD	STEM	ALLATIVE	GLOSS
kieli	kiele-	kiele-lle	'tongue'
käsi	käte-	käde-lle[9]	'hand'
hevonen[10]	hevose-	hevose-lle	'horse'
vene	venee-	venee-lle	'boat'
varis	varikse-	varikse-lle	'blackbird'
kirves	kirvee-	kirvee-lle	'ax'
johdin	johtime-	johtime-lle	'derivative'
askel	askele-	askele-lle	'step'
lyhyt	lyhye-	lyhye-lle	'short'

Likewise the comparative, superlative, and caritive morphemes have sandhi-determined allomorphy differences:

iso-mpi	iso-mpa-	iso-mma-lle[11]	'bigger'
iso-in	iso-impa-	iso-imma-lle[11]	'biggest'
raha-ton	raha-ttoma-	raha-ttoma-lle	'penniless, without money'

A morpheme conditioning stem allomorphy is an affix; a morpheme failing to condition stem allomorphy is a full word. Here the particle clitics pattern with the full words -- they attach to nominatives rather than stems: *kieli-hän (*kiele-hän)*. We have now strong evidence that these are words.

g. Ordering.

The ordering of a morpheme with respect to adjacent morphemes is indicative of the word/affix status of that morpheme. Freer ordering is typical of full words, e.g. (Hakulinen & Karlsson 1979:160-1)

(74) a. Leena korjasi laituria koko iltapäivän.
b. Laituria Leena korjasi koko iltapäivän.
c. Koko iltapäivän Leena korjasi laituria.
d. Koko iltapäivän laituria korjasi Leena.
e. Korjasi Leena laituria koko iltapäivän.
'Leena repaired the dock all evening.'

Strict ordering is typical of (especially inflectional) affixes:

(75) linnu-i-lle -- *linnu-lle-i 'to the birds'

The clitics show indeterminant behavior here. On the one hand, the particle clitics show strict ordering with respect to surrounding morphemes; on the other, they exhibit a slight degree of freedom as exemplified in 2.3.2 above. In a phrase such as the following NP the clitic *-hAn* may fall in either of two slots:

(76) a. nämä kauniit omenat-*han* 'these very pretty apples'
b. nämä-*hän* kauniit omenat 'ibid.'

In this otherwise free constituent order language the (rather) strict positioning of the clitics within the sentence stands out as atypical of syntactic words. In sum, then, the particle clitics are atypical affixes insofar as they show a certain degree of freedom and atypical as words in that they lack the full freedom of independent words. Not all independent words in Finnish exhibit syntactic freedom in ordering, as I point out in section 2.4 below, using several adverbs in the same semantic realm as the particle clitics. Thus, this test is inconclusive.

h. Distribution.

Typically the distribution of an affix is governed by a single principle. The illative, for example, combines with nominal words (adjectives and nouns -- including deverbal adjectives and deverbal nouns). But independent words rarely have distributions that are easily describable in a single principle -- the combinatory possibilities for a verb like *korjasi* are numerous. The particle clitics behave like affixes with respect to this criterion; they have simple statements that cover their distribution -- *-hAn*, *-pA*, and *-kO* combine with the first constituent of a clause, (sentential) *-kin/-kAAn* combines with the finite verb of a clause. This test constitutes only supportive evidence that the particle clitics are affixal. There are a number of monomorphemic words in the same semantic/pragmatic field as the particle clitics which have the exact same distribution. Thus this test cannot be used as conclusive evidence that the morphemes in question are affixes.

Complexity is introduced here when a pronoun occurs as part of a noun phrase, since WH-pronouns tend to attract clitics away from the end of the phrase (Östman 1977, Karttunen 1975c):

(77) a. missä maassa-han 'in which land'
b. missä-hän maassa 'idem.'

This special attraction on the part of the WH-pronoun is predictable and easily incorporated into our single statement of distribution; see section 2.2 above and chapter 4.2. But they are indicative of word status rather than affixal status.

i. Complexity.

Affixes are often morphologically simple, words are frequently morphologically complex. The particle clitics pattern with the affixes of the language in that they are morphological simplexes -- monomorphemic and non-decomposable. By contrast, the verb *korjasi* has not only a stem *korja(t)-*, but also the past tense *-(s)i* and an unmarked third person (cf. *korjasi-n* 'I repaired', *korjasi-t* 'you repaired', etc.)

Although the particle clitics are monomorphemic, the analysis of them as words is not incompatible with this test. There are numerous clear words in the language which are monomorphemic, including several in the same semantic field as the particle clitics (e.g. *muka* 'supposedly', *toki* 'truly', and the like). This test may provide supportive evidence of an affix analysis, but would not be strong evidence. This test is intended to provide strong evidence for a word analysis only if a polymorphemic element is examined. Otherwise, only supportive evidence is provided.

j. Syntactic Rules: Deletion.

Syntactic rules normally affect grammatical words (which are syntactic units), which may or may not happen to be equivalent to correspondent phonological words. Proper subparts of words, however, are immune to syntactic rules. This section and the next exploit this difference between words and affixes.

Zwicky tells us that

> Proper parts of words are not subject to deletion under identity; whole words may (in the appropriate circumstances) undergo such deletions (1984a:153)

For example, the full word *korjasi* may undergo deletion under identity (a,b), while the allative suffix may not (c,d).

(78) a. Leena korjasi laituria ja Seppo korjasi kelloa.
Leena fixed dock and Seppo fixed clock.
'Leena fixed the dock and Seppo fixed the clock.'
b. Leena korjasi laituria ja Seppo kelloa.
Leena fixed dock and Seppo clock.
'Leena fixed the dock and Seppo the clock.'
c. linnu-i-lle ja pu-i-lle
bird-PL-ALL and tree-PL-ALL
'to the birds and to the woods
d. *linnu-(i-) ja pu-i-lle
bird-PL and tree-PL-ALL
'to the birds and woods'

The particle clitics can undergo this deletion under identity, and perhaps **must** undergo it.

(79) ??auto-lla-han ja bussi-lla-han he tulivat.
car-ADES-HAN and bus -ADES-HAN they came
'By car and by bus they came, you know.'
(80) auto-lla ja bussi-lla-han he tulivat.
car-ADES and bus-ADES-HAN they came
'By car and by bus they came, you know.'

Sentence (79) is not necessarily ungrammatical according to my respondents, but seems to be unacceptable for other reasons. The repetition of pragmatic particles is in general disfavored, as in the sentences with *muka* below. This is as yet another way in which the particle clitics pattern with full words rather than with affixes.

(81) ??He tulivat *muka* bussilla ja *muka* autolla.
they came supposedly bus-ADES and supposedly car-ADES
'They came supposedly by bus and supposedly by car.'

(82) He tulivat *muka* bussilla ja autolla.
they came supposedly bus-ADES and car-ADES
'They came supposedly by bus and by car.'

At any rate we can say now that the particle clitics do not pattern with the affixes and do pattern with full words.

k. Syntactic Rules: Movement.

Affixes are not subject to 'movement rules' (in the terms of transformational grammar) and cannot serve as gaps in 'gap-filler' relations with other constituents in a sentence (in the terms of GPSG). Words such as *korjasi* can be moved to virtually any position in the sentence -- as in the topicalization in (74e) above. The illative, in contrast, cannot be moved without its host.

Taking for the moment a transformational approach, such as one found in Hakulinen and Karlsson (1979), we can cite several rules that suffice to show interaction between the particle clitics and syntactic movement rules. According to Hakulinen and Karlsson (p. 303ff), such rules include compensatory movements, heavy constituent postposing, and textual cohesion.

Let us examine only the first two of the three types of rules. As a compensatory movement rule, there is Compensatory Thematization, which functions to move a constituent following the verb into sentence-initial position in the instance that the (thematically neutral) verb would otherwise fall in that position. For example, in the sentence below the indefinite verb *asennetaan* 'one installs' does not take an overt subject (well, the subject is incorporated into the verb as the suffix *-ta-*). Because nothing else lies in sentence-initial position, some other (post-verbal) constituent is positioned there as the theme of the sentence (Hakulinen and Karlsson 1979:303):

(83) Asennetaan taloon sähkölämmitys
One installs house/ILL electric heating

---> Taloon asennetaan sähkölämmitys
house/ILL one installs electric heating

'One installs into the house electric heating'

In the "basic" word order there is a distinct avoidance of (textually neutral) verb-initial sentences in Finnish (Hakulinen 1984b), so that this rule would apply in most instances -- except when the verb is the only member of the sentence. Heavy Constituent Movement can interact with Compensatory Thematization, when the former applies to the theme. In the following the constituent *monet burleskit detaljit* is postposed due to the application of HCM, then Compensatory Thematicization fronts the NP *filmiä*.:

Keventävät filmiä monet burleskit detaljit.
lighten/3PL film many burlesque details

---) Filmiä keventävät monet burleskit detaljit.
film lighten/3PL many burlesque details

'Many burlesque details lighten a film.'

These various rules can move constituents in and out of sentence-initial position. But no matter what constituent gets placed or removed from initial position, the clitics *-hAn*, *-pA*, and/or *-kO* attach to the sentence-initial constituent. Thus the first constituent in a sentence may be a subject, a theme, a focus, a compensatory theme, etc.

All other affixes in the language follow the constituent to which they are attached. The particle clitics exhibit remarkable independence in this regard.

Summary.

In the chart below I summarize the results of tests a-k. The clitics pattern with the affixes with respect to certain of the tests, with independent words with respect to others.

PHONOLOGICAL TESTS:	
a. internal/external sandhi	word-like
b. prosodic phonology	affixal
c. word/phrase domain in segmental phonology	affixal
MORPHOLOGICAL TESTS:	
d. binding	affixal (supportive evidence)
e. closure	word-like (supportive evidence)
f. construction	word-like
g. ordering	word-like (inconclusive results)
SYNTACTIC TESTS:	
h. distribution	affixal (inconclusive results)
i. complexity	affixal (supportive evidence)
j. deletion	word-like
k. movement	word-like?

From all of this we can see that the particle clitics are not quite affixal, not quite word-like either. They seem to have an intermediate status between that of an inflectional affix and that of a full word. For this reason we shall continue to label these as 'clitics'. Later, however, I shall give up this label in favor of the **bound word/phrasal affix** division. Note, though, that the phonological tests point to an affixal analysis (albeit a loosely attached affix), whereas the morphological and syntactic tests indicate that a word analysis is more appropriate. The peculiarities that have emerged in the application of these tests will be addressed in chapter IV, when I confront my bound word analysis. The results here indicate we have the profile of a bound word--a separate word in syntax that is readjusted to form a phonological word with its neighbor.

In the following two chapters I examine the bound word behavior of the particle clitics. In chapter III, I present a fragment of Finnish syntax, with special reference to adverbial and conjunctive members of the same semantic (and, as I will argue, syntactic) class as that of the particle clitics. In chapter IV, I elaborate a bit on the bound word analysis of the particle clitics.

Endnotes.

1. Vilppula (1984) offers a few additional examples of sentential *-kin* where the clitic has none of the above functions. His corpus is so small that I do not take his analysis into account in this dissertation.

2. *-kA* is labelled 'lative' here because it attaches to pronouns inflected for directional cases: illative (e.g. *johonka* 'into which'), allative (e.g. *jolleka* 'onto which'), and translative (e.g. *joksika* '(becomes) which'). See Nevis 1984a for details.

3. By the terms **Wh-pronoun** and **Wh-phrase**, I mean the class of interrogative and relative pronouns and their accompanying phrases. These might also be labelled **Hi-words**, since all are related, at least etymologically, to the *mikä* paradigm. Hakulinen and Karlsson (1979:283) cite alternative terms for **Wh-question**, such as **hakukysymys** 'search question' and **täydennyskysymys** 'filler question'.

4. Apparently the source for Kaisse's Finnish facts is Lauri Karttunen.

5. Kaisse's data demonstrate a need for a taxonomy of conjunctions and connectives and their interaction with particle clitic attachment and with negative cliticization. The facts I offer here are representative of a possible approach to take. The attachment of particle clitics can in all likelihood be used as a means to distinguish between a veritable conjunction and a connective adverb.

6. My analysis accords with the intuition of my informant and colleague, Riitta Välimaa-Blum.

7. The clitic *-kin/-kAAn* is not a second position clitic, and is therefore not included with the other three here. This clitic attaches to the finite verb, which I assume is the head of the sentence.

8. Words like *toisen-kin-lainen* 'another kind, too' and *minkä-hän-lainen* 'what kind, I wonder' are apparent counterexamples, but a proper analysis of putative suffix *-lainen* 'kind, sort' shows that the morpheme in question is really a *cran*-type morph (i.e. a bound word in the lexicon). In these two intances, then, *toisen* 'another (GEN)' and *minkä* 'what (GEN)' are pronominal modifiers of bound noun *-lainen*. See also Karlsson 1983, Vesikansa 1977.

9. The *d* rather than *t* in *käde-ssä* is due to consonant gradation, a regular morphophonemic rule in Finnish.

10. *-nen* words also have *-s* allomorphs in compounds, e.g. *suomalais-ugrilainen* 'Finno-Ugric' -- cf. *suomalainen* 'Finnish, Finnic'. Strictly speaking, when I refer to stem allomorphy 'combination' here, I mean syntactic, rather than morphological combination.

11. *-mm-* in both the comparative and the superlative here (in place of *-mp-*) is the result of the regular morphophonemic rule of consonant gradation.

CHAPTER III

ON FINNISH SENTENTIAL ADVERBS

3.0. Introduction.

In the preceding chapter I noted the mixed status of the Finnish particle clitics, demonstrating that they pattern syntactically with full words, phonologically with proper subparts of words. I show in this chapter how the word-like properties of these clitics can be captured in the syntax of the language (by means of independently motivated syntactic operations). I begin with some general comments on the characteristics of sentential adverbials (section 3.1). I follow up with a section (3.2) that presents a framework to handle the syntax of the relevant adverbs. This chapter provides the background for a syntactic treatment of the syntactic parameters of the particle clitics (chapter 4).

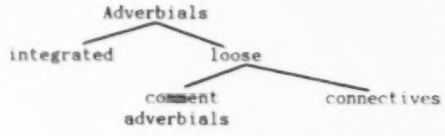
The particle clitics belong semantically to the class of **modal adverbs** (and the closely related set of **conjunctions**). The particle clitics, the modal adverbs, and the conjunctions do not interact with the propositional meaning or truth-value of the sentence; they contribute only additional information to the sentence and its truth conditions (Hakulinen and Karlsson 1979:255-6). This set of adverbs is semantically heterogeneous (Vuorinicmi 1973:283-4); that is, it has highly vague and diffuse semantics (Hakulinen 1984a:15). In addition, most members of this class are inherently *stressless*.

3.1. General Remarks on Finnish Sentential Adverbs.

I examine only the sentential functions of the adverbs in question, but note here that some of them can have scope smaller than the clause (Vuoriniemi 1973:283-4). For example, both full adverb *myös* and particle clitic *-kin/-kAAn*, meaning roughly 'also, even, too', have scope at different levels -- sentential scope and local scope (of an X' or X level).

I follow the taxonomy of adverbials given by Hakulinen and Karlsson (1979:202-6), who distinguish first between integrated and

loose adverbials, and second, in the latter group, between comment adverbials and connectives:



(Hakulinen and Karlsson 1979:202)

The integrated adverbials are sensitive to verbal properties and sentence type. The loose adverbials, by contrast, can appear in a sentence independently of what type of verb there is in the sentence and what sentence-type there is (Hakulinen and Karlsson 1979:203). Among the loose adverbials, then, 'comment adverbials' express the speaker's understanding of the proposition's degree of certainty, truthfulness, or origin, often also the speaker's emotive attitude to what she or he says. Comment adverbials generally only appear in declarative sentences, especially main clauses (or main clause-like, non-restrictive clauses of a type that could be considered subordinate clauses -- Hakulinen and Karlsson 1979:206). Examples of comment adverbials include *ehkä* 'maybe', *mahdollisesti* 'possibly', and *muka* 'supposedly'.

Connectives, on the other hand, function as linkers to sentences that came earlier in the discourse, and typically appear towards the beginning of a sentence (due to their semantics and pragmatics). In this regard, Hakulinen and Karlsson tell us, the connectives are closely related to the class of conjunctions. There are two main differences between connectives and conjunctions: conjunctions lie outside the sentence (i.e. they are daughters of S' rather than S), and connectives are open as a word class, whereas conjunctions form a closed word class. Examples of connectives are *taas* 'again', *kuitenkin* 'however', *lisäksi* 'in addition', *että* 'but, what if'; examples of conjunctions are *ja* 'and', *jos* 'if', *vaikka* 'although', and *koska* 'when, since, because'.

I treat the comment adverbials, connectives, and conjunctions as members of the same general class, marked diacritically with some arbitrary feature, say [class 41]. This [class 41] has subclasses, as in figure (1).

Among the subclasses of [41] are the comment adverbials [class 42] and connectives/conjunctions [class 43]. [class 43] is composed of two subclasses -- connectives [class 45] and conjunctions [class 44]. The class of conjunctions, [class 44], subcategorizes for S as a *sister* node. Connectives [class 45], on the other hand, form a subclass of [class 43] composed of adverbials *subordinate* to the S node.

I use arbitrary numbers here to refer to the relevant classes of adverbials in Finnish. This practice does not, ipso facto, curtail the role of universal properties in determining the membership within each class or subclass. I ignore here the issue of the naming of these classes and the role of universal semantic features in determining which adverbs may be members of these classes and subclasses.

Adverbial class 41 consists of both simplex adverbs and complex adverb phrases (as well as adpositional phrases):

Advl	Advl	Advl	Advl
|	|	|	|
Adv	Adv'	Adp'	V"

Here we are interested in the simple adverbs. These sentential adverbs (of class 41) cannot be modified and tend to occur toward the beginning of a sentence (i.e. sentence initial position or after the theme of the sentence). They are generally a main clause phenomenon (due to their semantics). Many members of class 41 occur in declarative sentences, but some are interrogative by nature. These include *entä 'what about'*, *kai 'perhaps'*, *kaiketi 'probably'*, *varmaan 'likely'*, *näinköhän 'this way?'*, *tokko -- tokkopa -- tokkohan 'hardly'*, *vai 'or'* and *-kO 'whether'* (Penttilä 1957:540). The interrogative adverbs seem to have a greater chance of appearing in subordinate clauses than other sentential adverbs; again this is a semantic fact rather than an arbitrary syntactic fact.

3.2. Fragment of Finnish Syntax.

In this section I present a fragment of Finnish syntax, couched in a modified generalized phrase structure grammar (GPSG) framework. I offer an account of the positioning of sentential elements, including the members of the verb phrase in Finnish. I demonstrate that at least some of the sentential adverbs of class [41] make crucial reference to the left margin of a sentence. My account of these sentential adverbs can then be extended easily and in a straightforward manner to handle the syntactic positioning of the particle clitics.

3.2.1. On Generalized Phrase Structure Grammar.

Of the several versions of GPSG available in the literature, I adopt that of Gazdar, Pullum, and Sag (1982) and Gazdar (1980),

incorporating however the modifications recommended by Gazdar (1985). Gazdar (1980) establishes the basic tenets of GPSG: two kinds of PS-rules, rule schemata, complex symbols, and feature conventions.

There are no transformations or coindexing devices. Gazdar argues that GPSG is a highly restrictive theory of generative grammar since it "is provably capable of generating only the context-free (CF) languages and is, to all intents and purposes, simply a variant of CF phrase structure grammar" (1980:2). He criticizes the continued use of movement and deletion rules, non-local filters, and/or coindexing devices on the grounds that such a framework is relatively unconstrained (or "at best one about whose constrainedness we know very little"). This highly constrained theoretical framework will allow us to make very strong universal claims about the properties of natural language.

GPSG assumes some version of X-bar grammar; in fact, Gazdar (1980) assumes a two-bar system. Syntactic constituents such as noun, noun phrase, and the like are represented as complex symbols. 'Noun' for instance is [+N,-V], though by convention I continue to use the cover symbol N and other familiar symbols in lieu of complex symbols like [+N,-V]. Features (located on the nodes) are used in abundance -- case features, verb types, classes and subclasses of syntactic word groups, etc. -- and are determined by admissibility principles on the syntactic nodes in a structure. Metarules are used to map rules into other rules (i.e. they establish relationships between rules -- of the sort 'If F is part of the grammar, then F' is also part of the grammar'). Finally, GPSG allows two (and only two) types of phrase structure rules: immediate dominance (ID) rules and linear precedence (LP) statements. ID rules establish syntactic configurations and assign semantic interpretations to constituents; LP statements give relative orderings for sister constituents. All ID rules precede all LP statements. I turn now to the division of rule-types in GPSG.

For a language like Finnish with a large amount of free constituent order in the sentence, but rigid word order in the phrases, it is necessary to keep immediate dominance separate from linear precedence. All constituents have configurations assigned by ID rules, but unless constituents must occur in a particular linear order, they need not have LP statements.

ID rules are formalized with three parts: first an arbitrarily selected rule number; second, the structural configuration (not unlike a traditional transformational rule) in which the mother category lies to the left of the arrow and daughter categories to the right; and third, a semantic interpretation statement (about which I will have nothing to say and therefore leave out):

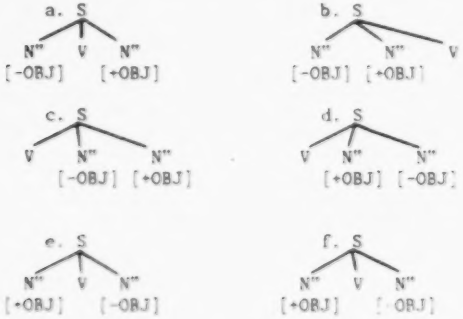
For example, Finnish permits all six orderings of the constituents in the sentence *Juha lyö Heikkiä* 'Juha hits Heikki'. Some orderings are clearly more favored than others; all are possible.

(1) a. **Juha lyö Heikkiä**
b. **Juha Heikkiä lyö**
c. **Lyö Juha Heikkiä**
d. **Lyö Heikkiä Juha**
e. **Heikkiä lyö Juha**
f. **Heikkiä Juha lyö**

On the one hand, we want to state here that all three constituents (N" ***Juha***, V ***lyö***, and N" ***Heikkiä***) are daughters of S (=V"). On the other hand, we do not need to stipulate in the configuration rule any (basic) ordering of the constituents. Hence I offer, tentatively, the following rule to handle the sentences in (1a-f): (Rule <1> will be replaced below by other, more adequate rules.)

<1; S ---> N"[-OBJ], V, N"[+OBJ]>

In this rule, <1> is the rule number, a diacritic that can be used elsewhere to refer to this rule. The daughter nodes N" [-obj], V, N" [+obj] are given in no particular order. Since there is no accompanying LP statement to restrict linear order, all six orders are permitted:



The structures above correspond to the example sentences given in (1). The symbol *S* is used in place of *V"* for ease of exposition. For all intents and purposes, every place that I use *S*, the reader can substitute *V"*. V, here, is intended as the head of the sentence.

It should be noted that the use of relational notions like [±OBJ] can be predicted from the semantic interpretation that has been left out here. If I included a semantic interpretation with rule <1>, I would not need to use the feature [±OBJ] above. My use of this 'feature' is purely expository and empty of any theoretical claims. It is used only in lieu

of the semantic interpretation left out of the rule. The general practice in GPSG is to use case marking, e.g. [±ACC]. But case marking is of little help here since in Finnish a subject may be case-marked either [NOMINATIVE] or [PARTITIVE] and an object marked for one of [GENITIVE], [ACCUSATIVE], [NOMINATIVE], or [PARTITIVE]. I will not, however, address the case marking facts in this thesis.

Constituents are assigned configurations in which there is a *head* (alongside, possibly, other sisters) within the constituent. The mother node is the same as the head daughter so long as a rule does not state otherwise (see Gazdar 1985 on default head daughter assignment). 'Free' features will then be identical for mother and head daughter within a constituent (again, unless otherwise specified). Rather than having the head daughter identical to the mother minus one bar-level, the notion 'bar level' is determined in the same way as other 'free' features, i.e. mother and daughter are identical in the default case. Rule ‹1› could be rewritten in the following manner:

‹1; V" ---› N"[-OBJ], H", N"[+OBJ]›

Here H" would be assigned the value V" since it is identical to its mother V", except for the specification of the zero bar level. I revise this rule below in section 3.2.3. where default feature assignment plays a more central role.

LP-statements: Although Finnish generally permits any ordering for daughter constituents of V" (and V'), lower level phrases have quite strict ordering requirements. This is the case for N"-s, Adj"-s, Adv"-s, and Adp'-s. As a simple example of how strict word order is to be achieved, I examine the Adpositional Phrase.

Finnish has a small class of prepositions and a large class of postpositions, with a handful of words that go both ways. The ID rule sets up the appropriate configuration:

‹2; Adp' ---› H", N"›

Then two LP statements establish word order for the pre- and postpositions, which are distinguished diacritically through arbitrary features: [class 18], say, for prepositions, [class 19] for postpositions. Our LP statements look like this:

Adposition < N"
[class 18]

N" < Adposition
[class 19]

Rule ‹2› in conjunct with the above LP statements gives us the following adpositional phrases:

[*ilman maitoa* 'without milk' has a partitive NP, whereas *maidon kera* 'with milk' has a genitive NP.]

Adpositions like *pitkin* 'along' appear to have dual membership:

pitkin siltaa -- *siltaa pitkin* 'along the bridge'

But actually, such 'dual membership' adpositions are members of neither class 18 nor class 19. As a result, no LP statement is relevant to these freer adpositions, and both orders are, then, predicted to occur by rule <2>.

3.2.2. Sentential Elements.

A typical sentence in Finnish may consist of more than a subject N", simple V, and object N". There may also be sentential and V' adverbials, as well as several finite and non-finite verbs. I revise rule <1> in the following manner:

<1'; S ---> N", H°, V*, N", Adv1*, Adv1*
[-OBJ][+FIN][-FIN][+OBJ] [VP] [SENT]

There are now several things to note about this rule. At this point, I leave out the verb phrase node V'. On the surface there is not much evidence for a V' constituent (Hakulinen and Karlsson 1979:228). This issue will be addressed more directly in section 3.2.3 below. For now, however, it suffices to point out that a verb (regardless of whether it is finite or infinitive) and its complement may easily be separated, as in (2). More on the infinitives below in section 3.2.3.

(2) a. Hän halusi sanoa sen. 'He wanted to say it.'
b. Sen hän halusi sanoa. ibid.
c. Sanoa hän sen halusi. ibid.

Since no LP statement accompanies <1>, constituent order remains free here.

I use the asterisk convention in rule <1'> as a device to allow any number of daughters of the same type, starting at zero. Thus we can have several non-finite verbs and/or several adverbials (of both V' and sentential scope):

(3) a. Hakijat *muka sentään* osaavat ruotsia.
'The applicants *supposedly for all that* know Swedish'.

b. Seuraava esiintyjä laulaa *todennäköisesti hyvin*.
'The following performer sings *apparently well*.'

(4) a. Pertti haluaa *lähteä juhlimaan*.
'Pertti wants *to go celebrating*.'

b. Lasten saattaa *täytyä jäädä* kesäksi kaupunkiin.
'The children happen *to have to remain* in the city for the summer.'

In (3a) both adverbials *muka* and *sentään* are sentential. In (3b) *todennäköisesti* is sentential and *hyvin* is V'. (4a) exhibits two infinitives *lähteä* and *juhlimaan* in addition to the finite verb *haluaa*. (4b) displays infinitives *täytyä* and *jäädä*.

The adverbials in rule <1'> may be interspersed among the other daughter nodes of S. In (5a) both adverbials are sentential adverbials; in (5b) *ilmeisesti* is a sentential adverbial (a comment adverbial in fact), *taitavasti* a verb phrase adverbial.

(5) a. *Suomessa* vastustetaan *nykyisin* tupakanpolttoa.
'*In Finland* one *nowadays* opposes tobacco smoking'.

b. *Ilmeisesti* isäntä leikkasi paistin *taitavasti*.
'*Apparently* the host cut the roast *skillfully*.'

3.2.3. On the Finnish 'Verb Phrase'.

The verb phrase (V') in Finnish is rather elusive. At first glance there is not much motivation for a V' node (Hakulinen and Karlsson 1979:225-28). In fact, we want to ignore it for the purposes of ID rule <1'> and free constituent order within the sentence. Yet certain facts demand a verb phrase treatment --government, substitution, deletion, and ellipsis (in the terminology of Hakulinen and Karlsson 1979).

First, object noun phrases and verb phrase adverbials can be governed by a specific verb or verb class. For instance, the verbs *ruveta* and *alkaa* have the same meaning 'to start, begin', but idiosyncratically govern different cases:

(6) a. ruveta + illative of III infinitive --
ruveta sata-ma-an 'to start to rain'

b. alkaa + I infinitive --
alkaa sata-a 'to start to rain'

Such valence still holds for the 'last' or 'bottom-most' verb in a verb chain. Thus, a verb like *tykätä*, which means 'to like' if it governs the elative (but 'to hold, consider' if it it takes a regular object NP), will still govern the elative in the following sentence, even though the 'higher' verb *haluta* 'to want' takes an object NP in the genitive, accusative or partitive:

(7) Haluaan tykätä siitä
I-want to-like it-EL
'I want to like it'

(cf. *Haluaan sitä/sen* 'I want it' with partitive *sitä* or genitive *sen*)

So, it is the lower verb *tykätä* that governs the complement rather than the higher verb *haluta*.

Second, Hakulinen and Karlsson (1979) posit three transformational rules that refer to the notion 'VP'. In spite of the non-transformational approach taken here, GPSG must have some correspondents to these transformational rules. Nonetheless I present the transformational account from Hakulinen and Karlsson without offering any GPSG replacement -- my purpose here is not so much to formalize the relevant rules as to motivate the VP constituent.

VP-substitution in Finnish replaces the verb phrase in the second of two conjoined sentences with the adverb *niin* positioned clause-initially and the clitic *-kin* attached to the subject of the second sentence (Hakulinen and Karlsson 1979:225-226). It corresponds to English *do so*:

(8) Marssivalmistelut jatkuivat pitkään, mutta niin neuvottelut-kin.
'The march preparations continued for a long time, but so did the negotiations.'

Verb Phrase Deletion and Verb Phrase Ellipsis[1] delete the verb phrase (except for the finite verb) in conjoined and subordinated sentences (Hakulinen and Karlsson 1979:226,323-6). In the following (9a) is the result of VP Deletion applied to (9b):

(9) a. Saat auttaa, jos osaat Ø.
You may help, if you know how Ø.

b Saat auttaa, jos osaat auttaa.
You may help, if you know how to help.

So, there is evidence both for and against the verb phrase constituent in Finnish. In the framework of GPSG we can handle both views. First I offer some verb phrase rules that are used for case marking and government facts and for the GPSG counterparts to VP Substitution, VP Deletion, VP Ellipsis, and VP Gapping. But then I

make use of a 'Universal Liberation Metarule' that makes all daughters of the verb phrase node daughters of the matrix S node.

There are some apparent verb phrases that fail to permit the free ordering discussed above, cf. (10) below. These VP's all involve an infinitive head (of which there are numerous possibilities for instantiation) rather than a finite verb. There are good reasons for assuming that these apparent counterexamples are actually not true verb phrases, but are instead full clauses (i.e they are not daughters of the matrix S, but of a subordinate S).

(10) hän vaikeni [sanottuaan sen].
PRO impressed PRO-having-said it
she/he impressed (us) as having said it.

*sen hän vaikeni sanottuaan.
*sanottuaan hän sen vaikeni.

One argument against a VP analysis here is based on case marking. I noted in Nevis (1981) that the case marking of object NP's in finite verb phrases differs from that in clausal verb phrases.

The daughters of these V'-s are, then, not also daughters of the matrix S, but are rather daughters of a lower S or V'. These verb phrases can be identifed from the nature of the infinitive that serves as the head. The first and third infinitives pattern together as in (7) above, but others (to be discussed below) pattern as in (10). This latter group not only forms a close-knit verb phrase, it can be demonstrated to have certain sentential characteristics (as regards object case marking, possibility of a genitival subject or the passive equivalent to a genitival subject). I henceforth assign to these infinitival verb phrases the structures in (11):

(11) a. Advl b. N" c. Adj"
| | |
S S S

These adverbial (nominal and adjectival) clauses include various non-finite passive constructions (passive inessive II infinitive and the past and non-past passive participles) as well as the two active participles (past and non-past), the translative I infinitive, and the two active II infinitives (instrumental and inessive). See, for example, Hakulinen and Karlsson on the assignment of Advl to these constructions.

Now, the daughters of a subordinated V' (as in 8-9 above) will not be sentential elements, and therefore will not be interspersed among the various other daughters of S. Daughters of subordinated V' can only interact with each other. Only sisters of the matrix clause exhibit freedom of ordering with respect to one another.

In sum, then, some V'-s are integrated into the sentence as daughters of S; others are not. Our general rule <3> captures the general verb phrase configuration (for the purposes of valence and other assignments of morpho-syntactic properties).

<3; V' --> H°, V' N", ADV'>
[-FIN]

In conjunction with rule <3> we need a rule that introduces object noun phrases and verbal adverbials as sisters of the bottom-most verb. Rule <4> takes care of this need.

<4; V' ---> H°, N", ADV'>

Rule <4> leaves out all reference to feature values, such as [+FIN], because this rule is so general as to encompass all kinds of V'-s. In other words, only this bottom-most V' contains the object N" and/or a V' adverbial. It does not matter whether the head is finite or nonfinite. My other V' rules do not introduce anything but verbs and verb phrases. so they are inherently without objects and adverbs.

I treat the Finnish verb phrase along the lines of the presentation by Gazdar, Pullum and Sag (1982). The verb phrase (V') has as its head, V. The symbol V here is a cover term for the [+V,-N] categories. Certain features may be associated with V or V' (or V''). Among these is the feature [±FIN]. The head of a [+FIN] V' crucially involves [MODE], and it can take [PERSON] and/or [NUMBER]. If [PASSIVE] is present, then the finite verb also take the feature [PASSIVE].

3.2.3.1. On Finiteness.

Gazdar, Pullum, and Sag (1982:597) make the [+FIN] verb *tensed*, but in Finnish tense (i.e. [±PAST]) and mode are separate from the person and number markings. Joseph (1983:ch. 2) argues that a single feature, such as *tense*, cannot be considered a universal defining property of the notion **finiteness**; instead, he argues that we need a multi-factor definition of a given notion or feature, and that overlapping factors are central to our identification of the finite verb.

In Finnish one can distinguish between a finite verb and a nonfinite verb on several grounds. First, the finite verb may not be case-marked, though infinitives in the language frequently take case affixes (e.g. *tule-ma-ssa* is the inessive of the MA-, or third, infinitive). Second, a finite verb may take a person and number marker of a particular ending class (i.e not 'possessive suffixes'), whereas infinitives and participles either lack a person/number mark entirely, or else have a person/number mark from a different ending class (i.e they permit 'possessive suffixes', which, by the way, cooccur with nominals).

Thus, in Finnish the MODE property (with which PASSIVE will always cooccur if it is present at all) is the determination of finiteness in Finnish. One problem for this approach is the nonequivalence of the location of MODE (and PASSIVE) and the location of the person/number marks. The problem only arises in one type of construction, namely negative verb phrases.

3.2.3.2. On the Negative Verb.

In Finnish, negation is instantiated by an idiosyncratic negative verb (verb stem in *e-*). This verb might be labelled an auxiliary verb because it requires a verb to follow it. It further has irregular syntax as well as irregular morphology. First, the negative verb lacks mode, except for imperative mode. The mode mark in a negative V' lies on the negative complement.

(12) e-n sano-ne
NEG-1SG say-POT
'I might not say'

(13) e-t sano-isi
NEG-2SG say-COND
'you would not say'

Second, since mode and tense are so tightly connected, the negative verb also rejects tense markers:

(14) ei sano-nut
NEG/3SG say-PAST
'did not say'

*ei-nut

But as the above examples show, the negative verb serves as the locus for the person/number markers in a V'.

I assume, with traditional accounts of GPSG, that the finite verb will host the subject agreement features. However, in Finnish this does not hold for the negative verb, so I offer a special rule below that will ensure the location of person and number markers on the negative verb, rather than on the finite verb.

(5, V' ---> V, H'
[NEG] [NEG] [-NEG]
[+FIN] [PERS] [-PERS]
[NUMB]

Rule (5) guarantees that when the negative V' receives its person and number agreement features from the subject, the person mark will be passed on to the negative auxiliary verb, rather than the finite verb. The number feature is still carried by the finite verb, as shown by (15-16).

(15) e-mme saano-nee-t
NEG-1PL say-PAST-PL
'we did not say'

(16) e-n saano-nut
NEG-1SG say-PAST/SG
'I did not say'

Such complicated rules as <5> above appear to falsify my claim that the finite verb is the one marked for MODE. Perhaps the negative verb is the finite verb. However, there is an independent reason to think that the complement to the negative in (15-16) is the finite verb. The finite verb serves as the host for the *-kin/-kAAn* clitic in its sentential use, cf. chapter 2.1.1.4.

(17) En tullut-kaan kotiin viime yönä.
NEG/1SG come/PAST-ALSO home last night
'I did not come home last night either.'

*En-kään tullut kotiin viime yönä.

3.2.3.3. On the Negative Imperative VP.

The Finnish imperative counts, in some sense, as a mode feature. It does not cooccur with potential or conditional morphemes. But for the purposes of negation, the imperative does not count. There is a special negative imperative verb (stem *äl-*), which takes a special complement V'. In the following I offer a rule to account for these facts:

<6; V' ---> H*, V'
[+IMP] [+IMP]
[+NEG] [+NEG]
[-FIN]
[PERS]
[NUMB]

The combination of features [+IMP, +NEG, -FIN] has a morphological realization that depends on the person and number marks on the head verb:

(18) älä tule-^{X} 'don't come (SG)'
äl-köön tul-ko 'let him/her not come, don't let her/him come'
äl-kää-mme tul-ko 'let's not come'
äl-kää tul-ko 'do not come (PL & FORMAL)'
äl-kööt tul-ko 'let them not come, don't let them come'

In these V'-s the complement verbs all exhibit either the second person singular morpheme /-^{X}/ (see section 2.3.a for the instantiation of this morphophone) or the non-second singular '-ko'.

3.2.3.4. On the Passive VP.

What is usually termed 'passive' in Finnish is in fact not a true passive. It does not promote an object NP to subject. Instead it indicates that the subject of the verb is a human agent, but is also indefinite (Hakulinen and Karlsson 1979:255ff). This approach accounts for the ability of intransitive verbs to undergo passivization, (19), and the inability of impersonal constructions to undergo it, (20).

(19) nuku-ttiin '(some)one slept'
sleep-PASS/PAST

(20) a. lunta sataa 'it is snowing'
snow precipitates

b. *sade-taan '(some)one is precipitating'
precipitate-PASS

I continue to use the notation **PASSIVE** in this dissertation, but not as an equivalent to any universal passive, rather as a cover term for the feature combination [+HUMAN, -SPECIFIED] (adopted from Hakulinen and Karlsson 1979:255).

The rule for the passive is the following:

<7; V' ---> V, V'
[PASS] [PASS]

This rule will operate in conjunction with the other V' rules given above, rules <3-5>.

3.2.3.5. On the Infinitive VP.

As for the infinitives, these can occur multiply embedded. The unmarked infinitive is the I infinitive (the TA-infinitive). If a verb is subcategorized for case, though, it must take the III infinitive. Hence I posit the following feature cooccurrence restriction (FCR).

FCR 1 V[CASE, -FIN] ⊃ [+III infinitive]

Here the 'feature' [CASE] is used (informally) to stand for a certain list of cases (such as [ILLATIVE], [ELATIVE], and so forth).

So for example, the verb *ruveta* 'to start' (from above) governs the illative case, and according to FCR 2, it must then take the III infinitive. FCR 2 makes the correct predictions: *ruveta sata-ma-an* 'to start to rain', where *sataman* 'to rain' is the illative of the third infinitive. The unmarked infinitive is [-III infinitive], the marked infinitive is [+III infinitive].

In this section I have offered several V' rules in order to capture valence properties and feature distribution. In this next section I show what role these rules play in Finnish grammar.

3.2.4. Phantom Categories.

Now I have introduced conflicting V' and S rules. Rule <1'> by-passes the V' and rules <3-7> elaborate various verb phrases. These latter rules presuppose the introduction of a V' node, as in <8> below:

<8; S ---> N", V', Advl*>

The V' rules <3-8> are used to predict relations among constituents and to predict actual rules of the language (such as valence and other subcategorization facts). My V' rules, then, are **phantom rules** (as suggested by Zwicky 1985c), which serve only to provide antecedents for another set of ID rules. Phantom rules are used to predict actual ID rules, without serving to admit branchings themselves. Following Zwicky's train of thought, structures assigned by rules such as <3-8> above are made 'flatter' by means of a single *universal liberation metarule* (ULM):

ULM: *IF* A ---> B,X *AND* B ---> Y *THEN* A ---> X,Y

(where A is any category, B any two-bar category other than S, and X and Y any sets of categories).

Zwicky's ULM permits us to assign correct subcategorization facts and semantic interpretations to V' units, yet allows us to have the correct dominance configurations as in <1'>. The phantom V' node is introduced by rule <8> to establish subject-predicate relations. Rules <3-7> bring about verb chains. V' might also be necessary to handle verb phrase gapping and ellipsis (Hakulinen and Karlsson 1979).

The ULM now predicts that daughters of V' will also be daughters of S. Both sentential and verb phrase adverbials will now be predicted to be daughters of S and both groups of adverbials can be interspersed among other daughters of S (the finite verbs, non-finite verbs, and subject and object noun phrases).

3.2.5. The Syntax of Adverbials.

At first glance it appears that sentential and verb phrase adverbials have the same freedom of occurrence that other sentential elements have. But free constituent order holds only for Adv"-s and Adp'-s. Individual adverbs that function as adverbials (i.e. [class 41], the 'loose' adverbials of section 3.1) have a tendency to be characterized by more strict word order. Of the numerous possible adverbial placements predicted by rule <1'>, members of this adverb class can be restricted to anywhere from one to four particular

sentential slots, while other members can have no restrictions at all. Strict ordering will be captured by idiosyncratic features in the lexicon some of which correspond to ID rules, some to LP statements.

In the following I examine the restrictions imposed on the adverbs of class [41]. These restrictions refer to just four slots in the sentence: sentence-initial slot, sentence-second slot, pre-verbal slot, and post-verbal slot. Some adverbs occur in only one slot (e.g. *myös* 'also' in post-verbal position); some occur in two or more of these slots. The adverbs, such as *että* 'but', that do not have these restrictions are not assigned the appropriate subclass features, and therefore will not participate in the rules presented below.

The slots themselves will be determined by the ID/LP format. They are stated in such a manner as to refer to constituents rather than individual words. This reference to constituents rather than individual words will eventually be directly relevant to my syntactic account of the syntax of the particle clitics.

The notion **second position** may mean different things for different languages. For Finnish it means the slot after the first constituent of the sentence. For a language like Bulgarian it means the slot after the first **word** of the sentence. Languages such as Luiseño permit both interpretations (Steele 1976, Kaisse 1985). In what follows I present examples of sentential position in second position; some of the examples may have initial constituents of only one word -- this is a fact about the average constituent length in my examples rather than a generalization about what constitutes the notion **first** in the sentence. Henceforth reference to the notion **(sentential) second position** shall always be taken to mean the slot after the first constituent of the sentence.

3.2.5.1. *Sitä.*

Let us now examine a few instances of strict positioning of modal adverbs. The adverb *sitä* has been described by A. Hakulinen (1975). It functions either as a marker of personal experience or else it indicates an emotional reading of the sentence or even a metaphor. *Sitä* is strictly ordered in sentence-initial position or after the first constituent of the sentence.

(21) Kyllä *sitä* voi mennä maata-kin, jos tuntuu,
certainly can go sleep-ALSO if feels
ettei enää jaksa
that-not longer hold out
'Certainly you can go lie down, too, if you feel that you can't hold out anymore'

(22) *Sitä* luulee, että kun menee naimisiin, niin taivas se nyt
thinks that when goes married then sky it now
aukeaa ja *sitä* ollaan onnellisia elämänsä loppuun asti
opens and one is happy life/GEN/3 end/ILL until
'You think that, when you get married, then the sky, it opens up and one is happy until the end of life'

The regular position for *sitä* is second position. It occurs sentence-initially only when the finite verb immediately follows it, as in both instances of *sitä* in example (22). When there are combinations of several class [41] adverbs occurring in initial and second positions, *sitä* will always be a second position adverb, never in initial position. Example (10) shows this: the adverb *kyllä* selects between either initial or second position (section 3.2.5.4); but in combination with (second position) *sitä* it must be initial. Since initial *sitä* must be followed by a verb, it curtails the appearance of another second position adverb: *sitä adverb verb.
[41] [+fin]

An example similar to *kyllä* + *sitä* can also be found in *sitä* + particle clitic *-hAn*. Particle clitic *-hAn* must occur in second position, but it cannot interrupt *sitä* plus a following verb, so **sitä-hän saattoi* is ruled out. Instead, the verb occurs in initial position, with *-hAn* attached, following by *sitä*: *saattoi-han sitä* 'it happened'. Such facts demonstrate that *-hAn* interacts with word order in the same way that a full adverb like *kyllä* interacts with word order.

In the two examples above, the initial positioning of *sitä* is evident. For clarification I give the clause boundaries '10' & 11'):

(21') [Kyllä *sitä* voi mennä maatakin [jos tuntuu [ettei enää jaksa]]]

(22') [*Sitä* luulee [että [kun mence naimisiin] [niin taivas se aukeaa] ja [*sitä* ollaan onnellisia elämänsä loppuun asti]]]

Hakulinen (1975) suggests that the reason *sitä* precedes a finite verb (under certain conditions) is that there is a tendency to avoid verb-initial sentences, because such sentences have a highly marked pragmatic function (for example, imperatives and 'neutral' yes-no questions begin with finite verbs). Apparently the avoidance of verb-initial constructions is stronger than the need to locate *sitä* in second position, so *sitä* gets 'bumped' out of second position into initial position.

3.2.5.2. *Muka* and *Toki*.

The adverbs *muka* 'supposedly' and *toki* 'truly' are always in the second position of the sentence in their sentential uses:

(23) Vuoressa *muka* asuu peikko.
mountain/INES supposedly lives troll
'In the mountain *supposedly* dwells a troll.'

(24) Sitten *muka* ruoka oikein hyvältä maistuu.
then supposedly food right good tastes
'Then *supposedly* food tastes really good.'

(25) Kerro *toki*, mitä siellä tapahtui!
tell truly what there happened
'*Do* tell what happened there!'

(26) Kun *toki* raatsit pihajäniksen tappaa!
As truly want pet-rabbit kill
'That you *really* want to kill the pet rabbit!'

3.2.5.3. *Vain*, *Totta*, and *Kai*.

Vain 'only, just', *totta* 'true', and *kai* 'maybe' (in their sentential uses) have somewhat fewer restrictions on their positioning. *Vain* occurs in second position or is placed before the verb (27); *kai* is usually sentence-initial or occurs in second position (28); and *totta* occurs clause-initially or before the non-finite verb (29).

(27) Onkohan *vain* oikein järkenevää panna rahojaan tuollaiseen.
is-Q-HAN only right frustrating put moneys such/ILL
'Isn't it just really frustrating to put one's money into something like that'

(28) Se *kai* sinun pitäisi tietää.
it maybe you should know
'THAT you should maybe know.'

(29) *Totta* hän meidät irti laskee.
true s/he us free lets
'No doubt he will let us free.'

3.2.5.4. *Kyllä* and *Edes*.

Kyllä 'certainly' and *edes* 'even' have somewhat more freedom than the above mentioned adverbs. *Kyllä* is generally sentence-initial (30) or else occurs in second position or before the verb (31).

(30) *Kyllä* hän on hyvin rikas.
certainly s/he is very rich
'He is certainly very rich.'

(31) Minä *kyllä* tulen.
I certainly come
'I will surely come.'

Kyllä will not occur in second position when there is an initial *sitä*, as I mentioned above (in 3.2.5.1). This is because *sitä* will be initial only if it is immediately followed by the finite verb.

Edes, on the other hand, can occur anywhere except clause-initially. This adverb does not occur in initial position because it is preempted by another word, namely the negative verb. Independent studies (Kangasmaa-Minn 1967) have shown that the negative has a strong tendency to occur sentence-initially. Since *edes* lacks the arbitrary restrictions found with the other adverbs of [class 41], it is left out from further discussion.

3.2.5.5. *Myös*.

Finally, let us consider *myös* 'also, too' (described by Östman 1977). In its sentential use it appears after the finite verb (Östman 1977:175):

(32) Hän on *myös* ostanut auton.
(s)he is also bought car
'She/he has also bought a car.'

As a variant of *myös*, there is also pleonastic *myöskin/myöskään*, with *myös* plus its clitic synonym *-kin/-kAAn*.

3.2.5.6. Summary of adverb positioning.

My data, then, exhibit four sentential slots, and various combinations thereof, in which these adverbs (that is, [class 41] occur (cf. chart below). Now, most of the members of class [41] have no real restrictions, other than tendencies for certain positions. So the default for this class is sans position stipulation. For the other members, each restriction is an idiosyncratic fact that must be listed somehow in the lexicon. I address exactly how this listing is to be achieved below in section 3.2.6., but first I explore the restrictions imposed on *kyllä, sitä, kai, totta, että, vain, muka, toki, and myös*.

	INITIAL	SECOND	PRE-VERB	POST-VERB
kyllä	+	+	+	-
sitä,kai	+	+	-	-
totta	+	-	+	-
entä	+	-	-	-
vain	-	+	+	-
muka,toki	-	+	-	-
myös	-	-	-	+

Myös is the only adverb, so far as I been able to discern, whose positioning makes reference to the postverbal position. Hence *myös* will have its own idiosyncratic mark in the lexicon. Among the other adverbs, some generalizations can be made. First, all occur in either initial or second position. In addition, the fact that some also occur in the pre-verbal slot is contingent on occurrence in initial or second position; I have found no members of class [41] that occur exclusively in pre-verbal position. This cooccurrence will be handled by an implicational statement.

To formalize the relevant restrictions, I establish the four parameters in arbitrarily labelled, binary lexical features:

[A] is the feature used for reference to initial position
[B] is " " " " " to second position
[C] is " " " " " to pre-verbal position
[D] is " " " " " to post-verbal position

As mentioned above, only *myös* carries feature [+D]. And any adverb having feature [+C] must also have either [+A] or [+B]. The following implicational statement handles this cross cutting feature:

(33) [+C] $\supset$ [+A] v [+B]

As for the adverbs *kyllä, sitä, kai, totta, entä, vain, muka, and toki*, these have the following features assignments, given in disjunctive statements. I have numbered the relevant subclasses, and listed members of each subclass which I have already discussed above.

41.1 kyllä (+A) v (+B) v (+C)
41.2 sitä,kai (+A) v (+B)
41.3 totta (+A) v (+C)
41.4 entä +A
41.5 vain (+B) v (+C)
41.6 muka,toki +B

As I stated above, these features are idiosyncratically attached to certain members of class [41]. Such idiosyncratic features do not belong in the syntactic component, as they do not serve any grammatical inflectional function. Rather, they are lexical features -- properties of items of the lexicon. The default values for other members of this class of adverbs are [-A,-B,-C]. Members

that are not marked with the plus-values of these features do not have the restrictive position that I have discussed already.

So far as I know, these lexical features are properties only of adverbs of class [41]. It might also be the case that imperative verbs (which, as mentioned above in passing, also occur in initial position) are assigned the feature value [+A], but for now I posit, tentatively, the following implicational statement:

(34) [A,B,C] ⊃ Adv
[CLASS 41]

Feature value [+A] corresponds to the syntactically distributed feature [FIRST] (see section 3.2.6). As yet another implicational statement, we have:

(35) [41] ⊃ [FIRST]
[+A]

This means that any member of class [41] having feature value [+A] will be marked [FIRST]. Such a member is *entä*. *Entä* is marked in the lexicon, [41.4]. Subclass 41.4 has the lexical feature [+A], which is, according to implicational statement (35), linked to the syntactic **margin feature** [FIRST]. I discuss the implications and implementation of the margin features below in section 3.2.6. For now, however, I point out that the distribution of such syntactic features is a matter for ID, so the location of the adverbs marked with [+A] (such as *entä*) in a sentence is achieved through ID rather than LP.

There is no principled way to predict the assignment of the feature [FIRST] to subclass [41.4]. It is instead an idiosyncratic fact that must be listed in the lexicon. In light of this, I make the feature [FIRST] an inherent property of this subclass, via feature [A]. In this way any sentence that contains an adverb marked with feature [A] (e.g. subclass [41.4]) cannot also contain another daughter with the feature [FIRST]. I explain this in more detail below in section 3.2.6.

By contrast the other features refer to positions assigned through LP statements. Adverb class [41.6] (e.g. *muka, tokî*), for instance, occurs after the first constituent of a sentence, but before all other constituents, and no other sentential slots are available. The feature [B] handles this. It will correspond to an LP statement which refers to the slot after the first constituent. An LP statement of this sort can be made in the following manner, using conventional GPSG formalism, but I will argue below that a slightly different mechanism should be utilized.

X < Adv < X
[FIRST] [B] [FIRST:Ø]

This LP rule tells us that an adverb of class [41.6] must follow any constituent marked [FIRST] and must precede anything marked [FIRST:Ø]. [FIRST:Ø] picks out those bundles of features that do not include [FIRST] specified for any value. Although this LP statement accounts for the facts in an adequate fashion, I introduce at this point another variant of the formalism, namely '<<' which stands for 'immediately precedes':

X	<<	Adv
[FIRST]		[B]

This rule produces the same effects as the conventional GSPG statement above. But I deviate from standard GPSG theory in my use of the notation '<<' to stand for the notion '*immediately* precedes'. In regular GPSG notation, LP statements refer only to 'follows' or 'precedes' without being able to specify immediately contiguity. But I shall argue below that the notion immediately precedence is crucially involved in the statements for the pre- and post-verb adverb classes. I shall also argue that the standard notation is not sufficient to replace '<<' with respect to the slots that refer to the finite verb. In this way I am forced to make a modest addition to the theory.

As an example of crucial adjacent positioning I repeat the sentential location of the adverb *myös*. It must follow the finite verb with no other material intervening. I give *myös* subclass number [41.8]. Adverb subclass [41.8] obligatorily takes feature [D].

V	<<	Adv
[+FIN]		[D]

A mirror-image LP statement holds for the slot that corresponds to feature [C]:

Adv	<<	V
[C]		[+FIN]

The reader may wonder why I use both arbitrarily assigned subclass numbers and arbitrarily labeled features that correspond to an ID rule locating the feature [FIRST], and the two LP statements above. The reason is that some adverb classes, such as [41.1-3, 41.5], may have disjoint statements for location -- including disjunctive ID/LP statements as well as disjunctive LP statements. GPSG does not allow disjunction between the autonomous parts of syntax, ID and LP. Disjunction is, however, permitted in the lexicon.

It is very simple to state sentential positioning for elements occurring in only one slot in a sentence (i.e subclasses 41.4, 41.6, 41.8). But those occurring optionally in more than one slot need disjunctive statements, which may refer to two different theoretical

mechanisms, such as an ID rule and an LP statement. This instance is represented by members of class [41.2], which occur either in sentence-initial position or second position.

For this reason I opt for a lexical treatment of the features, with the disjunctive (and implicational) statements listed in the lexicon, alongside other arbitrary facts. Furthermore, I have not listed the features [A,B,C,D] as lexical features of specific words, but as features characteristic of certain adverb subclasses. I have done this in order to capture the generalization that several words may pattern together with respect to their sentential location. For example, *sitä* and *kai* pattern together with regard to their positional restrictions. It would be peculiar to have to state separately the fact that *sitä* takes either [+A] or [+B] and that *kai* takes either [+A] or [+B]. Instead I have grouped these together into a single subclass. It is now easy to account for the possibility that speakers of the language are able to add to or take away from any particular subclass.

The seven LP statements above capture the relevant positioning properties of the class [41] adverbs. It was necessary for me to introduce some more formalism -- disjunctive statements are not used in standard accounts of GPSG. But this is, again, only a minor addition to the theory.

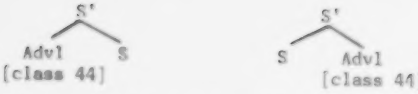
One can conceive of other position slots that could be relevant to a language such as Finnish -- for example, sentence final and penultimate positions, as well as combinations of the various slots. These slots apparently do not occur in Finnish, but there is no reason to exclude them from the set of grammars for other languages. It might turn out that the notion 'penultimate' is relevant in Finnish for a statement of the positioning of the conjunction *ja* 'and' (in a parallel manner to English *and* -- *ja* is, after all, a member of adverb class [41].

Instead, the absence of some of these LP statements that are 'missing' from Finnish can be attributed to pragmato-semantic facts -- modal elements have a nearly universal tendency to occur towards the beginning of a sentence (Steele 1975). Connectives and conjunctions also tend to appear toward the front of a sentence because of what they do.

Adverb [class 44] (i.e. the conjunctions, cf. section 3.1 , in particular, are daughters of S'. Rule ⟨9⟩ introduces this subset.

⟨9: S' ---⟩ S, Advl⟩
[class 44]

This rule predicts that conjunctions occur in only one of two positions:



Since conjunctions always precede S, we need an LP statement to rule out the righthand tree above

Advl < S
[class 44]

3.2.6. Margins in GPSG.

In the LP statements from the last section I introduce the features [FIRST] and [LAST] to refer to the margins of a constituent. These features are distributed as part of ID principles ; linking of features, such as [class 41.4] and [FIRST], or [class 41] and [LAST], is permitted. The margin features are 'percolated' from the mother node to one and only one daughter node. In some instances it does not matter which node that is so long as some constituent of S inherits it. If Adverb [class 41] inherits the feature, then conflict ensues from the LP statements referring to X[FIRST]. On the other hand, some members of adverb class [41] must (be permitted to) occur in sentence-initial position -- these are the adverbs that are idiosyncratically marked [FIRST] in the lexicon. The way this is done is through the linking of a particular subclass, say [41.4], with the margin feature: [FIRST:41.4].

Standard GPSG treatments do not permit reference to margins of constituents, so that this proposal of margin features is an addition to the theory. Reference to the margins of a constituent is clearly needed in syntax (cf. Zwicky 1985b), on a par with head and foot features.

3.2.7. Topicalization.

Some constituents which otherwise exhibit strict ordering of daughter elements, for example subordinated clauses, permit one (but only one) daughter to be fronted to the left of the matrix clause. This is known (in the literature on Finnish syntax) as 'topicalization'. Topicalization may apply to whole constituents such as N", Adp', V, or any elaboration of Advl, or it may go into one of these constituents to find a smaller constituent (such as an adjective, a noun phrase within a subordinated clause, etc. . In this subsection I examine this rule and its implications for the positioning of sentential adverbs in Finnish.

Although the daughter nodes of a subordinated sentence may not be interspersed among the daughter nodes of the matrix sentence (36b), it is possible to establish one (and only one) daughter of a subordinated clause in the initial slot of the matrix clause (36c).

(36) a. Mattilan täytyi myöntää Kertun nähneen Kallen.
'Mattila had to concede Kerttu's having seen Kalle.'

b. *Mattilan Kallen täytyi myöntää Kertun nähneen.
c. Kertun Mattilan täytyi myöntää nähneen Kallen.

This is the rule of 'topicalization', as discussed by Hakulinen and Karlsson (1979). This rule will front any constituent in a sentence (37), even out of a subordinated finite sentence (38):

(37) Hänet löysi yhteysalus Kumlinge.
him/her found community Kumlinge
'The community Kumlinge found him/her.'

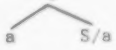
(38) ?Tällaisesta tiedän, että hän pitää.
this type I know that (s)he likes

To handle these topicalizations, I posit a rule of the sort presented by Gazdar, Pullum, and Sag (1982:602) and Pullum (1980):

<α; S --> a, S/a>

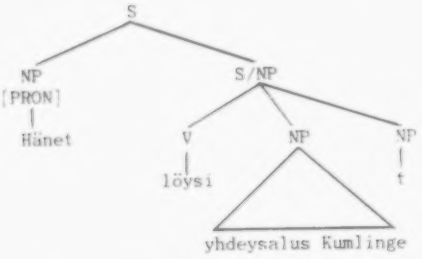
This 'slash' notation means that the S dominates a node a and a node S/a, and that S/a is a regular sentence with an a-type hole in it. Hence the structure in (39) is generated:

(39)



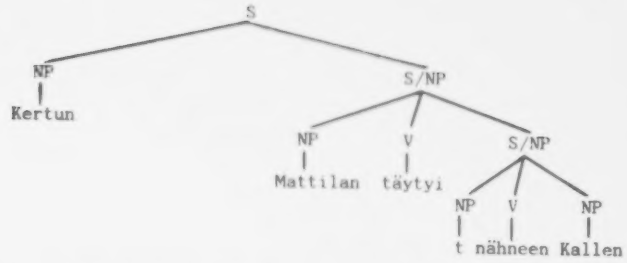
As a more concrete example I offer the following phrase structure tree for sentence (36) above:

(40)



In this structure the noun phrase *hänet* is topicalized, leaving behind a trace, *t*. A similar similar structure can be seen in (41):

(41)



Another example of topicalization involves the fronting of an adverb or adjective, with or without its complement NP. Karttunen (1975a:41-42) terms this 'focus', but it is the same as topicalization above:

(42) a. Kovin sinä olet vielä nuori.
quite you are still young
'You are still quite young.'

b. Sinä olet vielä kovin nuori.
'You are still quite young.'

(43) a. Uuden-han auton hän osti.
new-HAN car s/he bought
'It was a NEW car that he/she bought'.

b. Uuden auton-han hän osti
'It was a new CAR that he/she bought.'

In (42a) the adverb *kovin* is topicalized, yet the adjective *nuori* remains at the end of the sentence; in (42b) there is no topicalization. In (43a) the adjective *uuden* is topicalized; in (43b) it is the noun *auton* (or the entire noun phrase). A topicalized adjective or noun phrase will count as being first in its clause, as can be seen by *hAn* attachment in (43).

This holds true for WH-words as well. Since WH-pronouns have a special attraction for the particle clitic, as in (44), we know that they are marked as [FIRST] in their clauses. I assume that this marking is achieved in the same way as topicalization. WH-words, then, are inherent topics.

(44) missä-hän paikassa sinä asut.
what-HAN place/INES you live
'In what place do you live, I wonder.'

??missä paikassa-han sinä asut.

Sentential adverbs and particle clitics which refer crucially to the feature [FIRST] occur after the topicalized element rather than after the first daughter of the slashed S, as is the case with *-hAn* above. For this reason I assign the feature [FIRST] to the topicalized element in rule ‹9'›:

‹9'; S --> a, S/a›
[FIRST]

In this way the positioning of the adverbial element is stated correctly and the fronting of the topic is also captured -- the topic is marked [FIRST] and so does not appear linearly *after* the S-slash unit. This rule also prevents the assignment of the feature [FIRST] to any other daughter of S, since the margin feature is assigned to one and only one daughter of S.

3.3. Summary.

In this chapter I have set the scene for a syntactic approach to the positioning of the Finnish particle clitics. I have shown that, while the language is characterized by free constituent order, certain aspects of word order in the language (for example, within phrases) indicate that some linear ordering must be stipulated. I have established one class of sentential adverbs, class [41], which exhibit a tendency to appear sentence initially. Certain members of adverb class [41] show strict ordering; consequently I have posited subclasses here for the purposes of linear precedence statements.

I have also presented some basic immediate dominance rules to account for the flatter syntactic configuration of Finnish in comparison to that of a language such as English (Breckenridge and Hakulinen 1976). Finnish has no surface verb phrase, so my VP rules are used only to predict facts about government. The daughters of these VP nodes do not form a constituent of their own; rather, they are used to license other phrase structure rules. VP constituents get liberated into the sentence through Zwicky's Liberation Metarule.

As a consequence of the above rules, ID rules establish daughter nodes for a Finnish sentence (e.g. NP, V, infinitive, S-adverbial, VP-adverbial, etc.). Unless LP rules of the sort discussed in section 3.2.5 specify order, any constituent order may occur for the trees licensed by the ID rules. For members of adverb class [41], a complicated set of LP statements get the ordering correct. First, there is reference to the margin of the sentence (the left margin, in fact, in the guise of a feature [FIRST]). Second, the adverbs are positioned immediately before or immediately after a constituent (i.e. reference to the notion 'immediate adjacency'). And third, I had to introduce lexical features which refer to syntactic rules or statements, and which may have disjunctive statements.

I will point out in the next chapter that the syntactic positioning of particle clitics is so strikingly parallel to that of the members of adverb class [41] that repetition is introduced to a grammar of Finnish if both components -- syntax and cliticization -- permit exactly the same operations to apply. Instead, I shall propose to handle the syntactic aspect of the positioning of the particle clitics through the already motivated operations from the syntactic component. Duplication of function is especially questionable here because of the close semantic relationship between the particle clitics and adverb class [41]. In this way, syntactic operations will be purged from the cliticization component, and the component will form a more restrictive, more uniform component of grammar.

Endnote.

1. Hakulinen and Karlsson (1979:325) mention also VP-Gapping, which applies in the opposite direction of VP Deletion.

CHAPTER IV

TOWARD A SYNTACTIC ACCOUNT OF THE PARTICLE CLITICS

4.0. Introduction.

In this chapter I bring the material from the two preceding chapters together into a single, enlightening account of the syntactic behavior of the particle clitics. I show how the particle clitics pattern syntactically with the modal adverbs, connective adverbs, and conjunctions (i.e. adverb class [41] from section 3.1). One crucial aspect of this grammatical class is the strong tendency for its members to occur without any phrasal prominence whatsoever (though most permit emphatic or contrastive stress). These adverbs are generally without phrasal accent, and at least one member of this adverb class is inherently stressless. I point out that this unstressable adverb forms a link between the independent adverbs of class [41] and the so-called particle clitics. This will entail an examination of stresslessness with respect to a broader theory of prosody. I do not, however, establish any working theory of a component of prosody, although I do discuss the basic units needed in such a theory.

In the first section of the chapter (section 4.1), I examine the basic units of a component of sentence prosody, focussing on the difference between the **phonological word** and the **phonological phrase**. In section 4.2, I make explicit my view that the particle clitics of Finnish are really **bound words**, by demonstrating how the particle clitics fit (syntactically and prosodically) into the class of sentential adverbs and conjunctions. And in the final section of the chapter (4.3), I contrast the bound word type of clitic with another type of clitic, the phrasal affix.

4.1. Assignment of Phrasal Prominence.

In this section I explore the relationship between clitichood and stresslessness. I show that stress and accent are not definitional for clitic status, though there may be some correlation between prosody and clisis. Later in this chapter I drop the terms 'clitic' and 'cliticization' in favor of more appropriate

appellations, bound word and phrasal affix in place of clitic, liaison in place of cliticization.

The clitic is notoriously difficult to define in any strict sense. Investigators of particular languages and/or prosodic phenomena have often defined clitics as inherently stressless items. Selkirk (1984), for instance, tells us that the term 'clitic' is used to designate a word that "is stressless and immediately adjacent, juncturally or rhythmically speaking, to what follows or what precedes" (p. 340). Likewise, Lehiste (1982a:4) speaks of "lexical items (usually function words) hav[ing] no inherent stress, so that when they appear in a spoken sentence, they get attached to a neighboring word that does carry a certain degree of stress". Contra Selkirk and Lehiste (or, at least, a strict interpretation of their statements), I will argue in the following that stresslessness is not a sufficient criterion for establishing clitichood.

For reasons to be discussed below, I adopt a different line from that of Lehiste and Selkirk. It is true that a number of grammatical function words do not bear stress or accent as a matter of principle. But these can be distinguished from true clitics, which have additional properties. Zwicky (1985a), in particular, emphasizes that accent is susceptible to attack when used as a principle for determining clitichood. Klavans (1982a) comes to the same conclusion, and, in fact, cites a few examples of stressed clitics from the literature on clisis. Some of these examples involve word-level stress, rather than phrase-level stress, but the point is well taken anyway.

4.1.1. Stresslessness And Clitichood.

As an unproblematic example of the phenomenon of stressed clitics, let us consider Klavans' (1982a:104-5) discussion of Dieter Wanner's material on Southern Italian (Tyrrhenian dialect), in which clitics may accept word-level stress under certain conditions. The regular location of main stress, on the verb host, is then the locus for secondary stress or may even remain without stress entirely. These stressed clitics can be seen in the following examples taken from Klavans (1982a:104).

(1) (r)a-mé-nne -- (r)á-me-na 'Give me (some) of it!'
give-me-it

(2) tòrna-mé-llo 'Give it back to me!'
return-me-it

(3)ràten-gé-lle 'Give us (some) of it!'
give-us-it

(4) sposàre-sé-lla 'to get married to her'
marry-REFL-her

In the above examples, the clitics *-(m)me*, *-ge*, and *-se* bear main stress.

Klavans offers instances of stressed clitics also in French and Hixkaryana. Her discussions of English pronouns (e.g. *him* vs. *'m*) and Ancient Greek ("Enclitic Accent Throwback") are not particularly convincing, since one could easily make the case that these are word/clitic pairs (i.e. suppletive allomorphs) rather than an instance of restressing. Of importance here, though, is Klavans' claim that stressed clitics in French and Hixkaryana (from Derbyshire 1979) result from intonation rules that apply to whole phrases (including sentences). The French case is dubious on other grounds -- mainly the problem of establishing the notion word. The Hixkaryana example is, however, more convincing, as it involves two phrasal intonational rules, non-terminal and terminal. Both rules permit enclitics (or according to Derbyshire (1979), word-final bound words) to take intonational stress that is assigned to the final syllable of a phrase.

Klavans also has a third set of examples -- those that permit stress under conditions of emphasis or contrast (what she calls 'semantic stress'). For these she cites Turkish and Old Spanish.

Thus, there are examples of stressed clitics (although not many). These are always due to a conflict of principles: some general rule of prosody (Klavans discusses word-level accentuation, phrase-level accentuation, and emphatic/contrastive stress) assigns prominence to a particular syllable, and overrides the absence of main word stress on a particular morpheme or word. But conflict is usually avoided in most languages, where, in lieu of stressed clitics, one finds corresponding "full" or "strong" forms. Conflict avoidance is especially prevalent for those languages having pronominal clitics, like Serbo-Croatian. According to Browne (1967), Serbo-Croatian exhibits complementary distribution between full and clitic (second position) pronouns. Full pronouns are called for when there is contrast or emphasis; clitic pronouns otherwise.

(5) a. Da ti dam knjigu?
CONJ 2SG give book
'Should I give you the book?'

b. Da dam knjigu tebi?
CONJ give book 2SG
'Should I give YOU the book?'

In (5a) the pronoun *ti* is a second position clitic; in (5b) it is a full dative pronoun.

It is important to keep in mind the observation by Kaisse (1985), who notes that, although clitics may accept other kinds of stress, they do not bear a stress independent of the word on which they depend. From Klavans' description above, we can add: "... or independent of the phrase in which they occur."

The relationship between clisis and stress will not be so direct as Lehiste, Selkirk, and others indicate. Besides the existence of stressed clitics, there are also reasons to want to distinguish between true clitics on the one hand, and other unstressed words on the other. Zwicky (1982a) makes the point that clitics are characterized by idiosyncrasies not predictable from their stresslessness, properties such as special syntax, special morphology, or special phonology. Kaisse (1985) likewise makes the point of distinguishing clitic versions of English words from their merely unstressed counterparts. In particular, Kaisse notes that there are, arguably, phonologically reduced forms of *is*, *has*, *don't*, and *who* which are not the result of cliticization, but which can be derived solely by normal phonological rules of the language. Other allomorphs (such as [s,z, əz] for *is* and *has*, [dō] for *don't*, and [wə] for *won't*) are not so derived, and are marked in the lexicon. In fact, a main point from Klavans' (1982a) dissertation is the use of a feature [± clitic] in the lexicon to mark lexical items, since clitichood simply cannot be predicted from stresslessness or other factors. These facts have also led Zwicky (1982a) to predict stresslessness from clitichood rather than the other way around (a suggestion he attributes to Selkirk 1972). To combine Klavans' and Zwicky's viewpoints, then: [+clitic] implies [-stress].

4.1.2. Prosodic Units.

The discussion in the preceding section reveals that clitics are a subclass of stressless words (though see 4.2.2 and chapter 7). Clitics are inherently stressless (as are a number of grammatical function words), but clitics bear other properties as well. Such a view of the relation between clisis and stress is exactly parallel to a distinction drawn by Sadeniemi (1949) and Zwicky (1982a), among others[1], between the phonological word and the phonological phrase.

Sadeniemi (1949) establishes some rudimentary units of prosody: *the syllable, the measure, the phonological phrase*, etc. The phonological phrase (Finnish *'puhetahti'*; Lehiste (1970:164) translates this as 'speech measure') is composed of a syllable with main stress followed by syllables lacking main stress. Within the phonological phrase are sequences of 'measures' (Finnish *'tahti'*) which are composed of a stressed syllable (primary or secondary stress) plus any subsequent unstressed syllables. Sadeniemi (1949) makes a point of including unstressed

conjunctions, modal adverbs, pronouns, and other grammatical function words as part of the phonological phrase[2]. They attach prosodically to the nearest main-stressed syllable, if they do not become restressed for discourse reasons.

Zwicky (1982a) likewise distinguishes between the phonological word and the phonological phrase. His interest lies in demonstrating the non-equivalence of notions such as word and phrase at different levels of grammar. Zwicky discusses four uses for the notion **phrase**: the syntactic phrase (i.e. constituent), the semantic phrase, the lexical phrase, and the phonological phrase. Of importance to us here is the last in the group. Zwicky tells us that:

> **These phonological phrases** [emphasis original] serve as the domains within which external sandhi rules operate, and their boundaries are locations for the operation of phonological rules conditioned by hiatus or pause; they also serve as the domains for the assignment of stress and intonation patterns, and as units of timing; and their boundaries mark locations where parenthetical interruptions can occur. It seems unlikely that a single type of construct will be able to serve all functions at once, but in at least some simple cases we should be able to expect that the various criteria demarcating phonological phrases will coincide. (Zwicky 1982a:1)

Zwicky also assumes a principle whereby phonological phrases (and words) are syntactic constituents (and words), unless otherwise specified. What interest us here are the "principles for reorganizing and reducing syntactic structures, to yield the divisions appropriate for phonology" (Zwicky 1982a:3). The prosodic principles do not always provide a unique prosodic organization for what may otherwise be a unique syntactic structure (cf. Zwicky 1982b and Selkirk 1984); they may merely provide alternative reorganizations. In a similar line of work, Lehiste (1982a,b,c; 1983) assumes a mutual independence of syntactic and phonetic structures, with some sort of "code" to connect the two. Lehiste (1982b:123) reports, "the results obtained in this [phonetic] study show that in a large number of cases, the suprasegmental realization of a sentence cannot be predicted from its grammatical structure."

I follow very closely Zwicky's, Lehiste's, and Sadeniemi's viewpoints in these matters. I shall restrict the discussion, however, to the notions 'phonological word' and 'phonological phrase' in Finnish, and to the lack of complete isomorphy of these phonological-prosodic units with corresponding units of syntactic structure.

4.1.2.1. The Phonological Phrase.

The Finnish phonological phrase can be signalled in several ways: It is characterized by one main stressed syllable in conjunction with one or more other, non-main-stressed syllables. Stress itself is signalled by duration, pitch, and intensity. Niemi (1984) reports that stress, for word-level stress at least, is marked by a somewhat longer segmental duration in the prominent syllable than in a non-prominent syllable (both as regards the first vowel:second vowel ratio and a more complex measurement for 'timing precision'[3]), by a F_0 peak on the vowel nucleus of the stressed syllable, and, with a comparatively decreased role, by amplitude. The three parameters may vary in the role they play for individual speakers. For some F_0 is perceptually more important; for others, duration has the greater load.

These factors play a similar role in English, where further assistance in demarcating a phonological phrase is provided by juncture. Lehiste (1982a,c) reports that preboundary lengthening (as well as laryngealization, pause length, and postboundary lengthening) is characteristic of English phonological phrases as a signal of juncture, and may well play a bigger role than fundamental frequency. Fujisaki and Lehiste (1982) find similar facts for Estonian speakers as well, but Niemi (1984:34ff.) denies that particular rhythmic parameter for Finnish speakers. Instead, juncture may be signalled in Finnish phrases by "heightened precision" at the boundary of a nominal phrase-internal boundary (Niemi 1984:75).

4.1.2.2. The Phonological Word.

The phonological word in Finnish can be delimited on a number of criteria. Several morphological and phonological rules have, as their domains of application, the phonological word. First, there is Vowel Harmony. Back/front harmony applies within the phonological word, but not across (phonological) words:

(6) työ alkaa -- *työ älkää
work begins

It is limited to the phonological word and does not extend to the phonological phrase, as can be seen in (7) and (8). (7) has a particle clitic *-hAn*, which does undergo Vowel Harmony; (8) has an unstressed adverb *sitä*, which does **not** undergo VH.

(7) tuuli-han -- *tuuli-hän 'wind'
tyyli-hän -- *tyyli-han 'style'

(8) saattoi-han s(i)tä -- *s(i)ta
happened-HAN SITÄ
'It happened'

Vowel Harmony, however, functions as a juncture indicator only when there is a change in front/back harmony from one word to another. It does not act as an indicator of juncture when both words have the same harmony. For example, the phrase *pian on* 'soon is', with two phonological words, is not distinguishable from single word *pianon* 'of the piano' on the criterion of VH.

Another juncture indicator is the automatic insertion of aphonematic glottalization (either laryngeal restriction [ʾ] or glottal stop [ʔ]) before a word-initial vowel /avata/ ---> [ʾavata] -- [ʔavata] 'to open' (Karlsson 1983:133). This clearly serves to mark the beginning of a word. In connected speech this aphonematic segment may assimilate to a preceding consonant or may disappear altogether: /saat avata/ ---> [saatʾavata] -- [saat(:)avata] 'you may open'.

Yet another juncture-signalling rule is the assignment of word-level stress. In Finnish, main stress is assigned to the initial syllable of a (phonological) word, and secondary stress alternates every other syllable thereafter, as in (9).

(9) kǽytännȍllisèssä 'practical (INES)'

A main stress thus serves to mark the beginning of a word.

Karlsson (1983:150ff) mentions a few other means to indicate (phonological) word boundaries -- shortening of two consonants after a long vowel or diphthong in word-initial position; lengthening of a second syllable short vowel after a short syllable; and lengthening of a short consonant after an initial short vowel and before a long vowel or diphthong. In many instances phonotactics serve the same function (see Karlsson 1983, and references therein).

4.1.2.3. The Finnish Phonological Phrase: Some Problems.

The phonological phrase in Finnish can be recognized with some certainty, and has been described since at least Sadeniemi's time (1949). But the details of his analysis have yet to be explicated. For instance, Sadeniemi does not answer the question he raises concerning the determinants of the direction of prosodic leaning.[4] In (10) the pronoun *minä* leans left to the conjunction *mutta* rather than to the verb *rohkenen*. In (11) the pronoun *te* leans rightward onto the following main verb *olette*.

(10) mùtta-mínä róhkenen kúitenkin árvella ...
but I dare however to guess ...

(11) jòiden váikutùksesta té-òlette jóutuneet tämän máan jättämään.
REL.PRO. influence-EL you are ended up this land to leave
'under the influence of which you have ended up leaving this land.'

[Both examples from Sadeniemi 1949:53, with the addition of some standardization. Note also the restressing of the pronouns.]

A second question raised by Sadeniemi (1949:56) is the distinction between the attachment of main words to each other and the attachment of the stressless words to main words. He posits two different juncture symbols: a hyphen and three dots. In (12) and (13) the two kinds of juncture link main words together. Example (12) has *koettakaa* linked with *siis*. Example (13) has *elämän* linked with *vaiheita*.

(12) kóetta- -kàa- -sìis [oivaltaa minun äänestä harras myötätunto]
try IMP(2PL) then
'try then [to grasp from my voice the deep sympathy]

(13) énhän-mìnä túnne tèidän-jókaisen élämän...váiheita-ènkä
not-HAN I know your(PL) every life vicissitudes-nor
kókemùksia
experiences
'I do not know the vicissitudes and experiences of each of your life.'

[Again, standardized somewhat.]

Under Sadeniemi's treatment, many of the grammatical function words, such as pronouns, conjunctions, and certain adverbs, will never (or will typically not) be the locus of the main stress of the phonological phrase. Instead, these satellite words remain prosodically subordinate to the main stressed syllable of the phrase.

4.1.3. Leaners and Clitics.

In the following I demonstrate that a systematic difference between clitics and prosodic leaners can be made for Finnish.

Sadeniemi (1949:81) points out that a potential difference between purely grammatical and purely lexical occurrences of the same word might crop up. He gives, as an example, the word *juuri* 'just'. In its grammatical function it is modal adverb (i.e. a member of class [41] that signals emphasis) and therefore lacks phrasal accentuation, as in (14). In its use as a 'material adverb' it does not act as a grammatical function word, and permits a certain degree of stress, as in (15).

(14) Juuri hän oli täällä.
just (s)he was here
'Precisely she/he was here.'

(15) Júuri hän oli täällä.
'He was just here.'

Satellite words of the sort discussed by Sadeniemi are prosodically dependent on their stressed hosts, but they often may also be restressed (as in the case of examples (9-10) cited above).

Such prosodically dependent words are termed 'leaners' by Zwicky (1982a). Words that optionally lean on another prosodically may be called **optional leaners**. In addition, there are prosodic leaners that must lean prosodically on a neighbor. These might, then, be labelled **obligatory leaners** in Zwicky's terminology, or, according to Wackernagel (1892), **quasi-** or **semi-clitics**.

The adverb *sitä* is precisely one such semi-clitic. As a member of class [41] (especially subclass [41.2]), it exhibits syntactic independence, but is never stressed (it does not allow sentential stress of any sort — not even emphatic stress). Generally it loses its first vowel, *sitä* > *stä*, even if it leans to the right (e.g. *stä-ollaan* 'one is'). The resultant word-initial consonant-cluster is peculiar on phonotactic grounds, since Finnish lacks word-initial consonant-clusters (in its native inventory). Hakulinen (1975:27) further notes that *sitä* cannot stand alone in answer to a question.

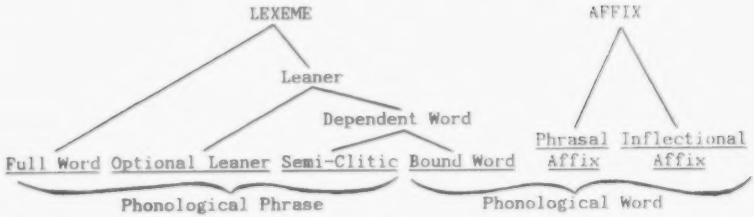
(16) a. Täällä sitä viihtyy.
here SITÄ gets along
'One will get along here.'

b. Kuka viihtyy täällä?
Who gets along here
'Who will get along here?'

*Sitä.

Optional leaners and semi-clitics are not to be confused with true clitics. True clitics exhibit *phonological* interaction with their hosts, whereas leaners exhibit only prosodic interaction. In the case of semi-clitic *sitä* and the particle clitics, this difference lies in the absence or presence (respectively) of the word-internal phonological rule of Vowel Harmony. *Sitä* is not subject to Vowel Harmony in Standard Finnish, whereas true clitics like *-hAn, -pA, -kO, -kAAn* are subject to this rule. Veritable clitics may then be viewed as a subset of leaners since they just like to leaners (in their syntactic and prosodic parameters), only with the addition of

phonological interaction with the host. (See also Nevis 1985b for a similar division of labor for clitics in Estonian.)



Many works that purport to deal with *clitics* actually treat *semi-clitics* (or even *optional leaners*). Such is probably the case with Russian interrogative *-I'i* which, so far as I can tell, cannot be proved to interact phonologically with its host. Otherwise it behaves exactly like a second position clitic would behave in a language. Along with Zwicky (1982a) and Kaisse (1985), I propose that any serious work on clisis must take pains to draw the division between clitics on the one hand and leaners (of both types) on the other. This will entail a reevaluation of all the phenomena that have been claimed to illustrate Wackernagel's Law, for instance. Only a few studies make this important distinction between leaners and clitics. Kaisse (1985) is one, who in her discussion of English *to*-contraction, Auxiliary Reduction, *don't*-cliticization, and *who*-cliticization, makes a concerted effort to keep the clitics separate from the leaners.

4.1.3.1. Additional Leaners and Clitics in Finnish.

My model of module interaction will also permit morphemes and words to switch fairly easily from semi-clitic to true clitic through time. The immediate historical predecessor to interrogative clitic *-kO* is semi-clitic *ko*. Invariant *ko* still exists in certain Finnish dialects (Sadeniemi 1949:43-44, L. Hakulinen 1979:88) and in Karelian (Laanest 1982) as an marker of interrogativity. The crucial difference between clitic *-kO* of (Standard) Finnish and semi-clitic *ko* of Finnish dialects and Karelian is the presence or absence, respectively, of phonological interaction with the host, in the guise of Vowel Harmony:

(17) a. clitic (Standard Finnish) hän-kö 'he/she?'
b. semi-clitic (dialectal) hän-ko 'he/she?'

Similarly, in some Finnish dialects, the semi-clitic *sitä* appears cliticized as *-stA*, pronominal leaner *mä* (from full pronoun *minä* 'I') as *-mA*, and conjunction leaner *ja* 'and' as *-jA*. I have already discussed the former two dialectal

clitics above in chapter 2. The last can be seen in (18) where the quality of the vowel in the conjunction[5] *jA* is determined by the host *mies* or *nainen*. *Mies* conditions front-vowel harmony, as in *mies-hän* not **mies-han*, so the vowel assigned to *jA* is front vowel *ä*. *Nainen* has a back vowel, *a*, so *jA* is realized as *ja*.

(18) mies-jä vaimo-ja lapsi (Beronka 1922:89)
man and wife and child

4.1.4. On Prosodic Theory.

Let me now return to the discussion of prosodic theory. It is true that the division between members of a major word class and those that belong to a closed, grammatical function-word class has been long recognized, as well as the overall association of (at least the availability of) stress with major word classes and lack of stress with minor word classes. The celebrated rule of English Nuclear Stress Assignment describes these facts to a large extent, but it remains controversial, and moreover, its applicability beyond English is dubious at best.

In the model of grammar that I present in chapter 7, I do not connect cliticization directly with prosody. Instead, I locate the two in adjacent components. Cliticization lies in a module of grammar whose function is to form units of connected speech (in the case of cliticization, **phonological words**). Prosody constitutes its own component, some of the purposes of which are to establish phrasal stress and to build **phonological phrases**. These two components, cliticization and prosody, share at least the property of being located in the grammar after the syntactic component.

4.2. Argument for Particle Clitics as Words.

In chapter 2 I demonstrated the mixed status of the particle clitics: on the basis of Zwicky's (1984a) diagnostic tests, the particle clitics have the syntactic properties of full words, but the phonological properties of affixes. In fact, if it were not for the bound status of the morphemes as well as their phonological interaction with the host, the particle clitics might well be considered regular words. This is exactly the viewpoint that I take: the particle clitics are syntactic words which happen to be phonologically bound.

Particle clitics *-hAn, -pA, -kO* pattern prosodically, syntactically and semantically with adverb subclass [41.6], and particle clitic *-kin/-kAAn* (in its sentential use) with subclass [41.8] (i.e. sentential *myös*). On semantic grounds both groups of morphs share vague semantics which do not interact with the truth conditions of a sentence. On prosodic grounds they pattern alike:

class [41] adverbs are generally stressless -- they are prosodic leaners, usually optional, sometimes obligatory. And on syntactic grounds the particle clitics and class [41] adverbs pattern alike in terms of the restricted number of sentential slots in which they can appear.

At the beginning of chapter 3, I discussed connective and modal adverbs as well as conjunctions. This set of grammatical function words is characterized by the following properties:

a) vague and heterogeneous semantics; usually a discourse, interpersonal, and/or textual function.

b) tendency to appear in certain sentential slots (especially initial position, second position, before the verb, or after the verb), and in some instances these tendencies become absolute restrictions on occurrence.

c) stresslessness -- either as a general tendency, or as an inherent feature of the word (as with *sitã*).

d) strong tendency to become prosodically attached to a neighbor, as 'leaners', but not phonologically attached.

A brief comparison reveals that the particle clitics are characterized by the exact same properties:

a') vague and heterogeneous semantics, with a clearly textual and/or interpersonal function.

b') strict positioning in the sentence: either in second position, or after the finite verb.

c') inherent stresslessness.

d') phonological attachment to the last word of a preceding constituent.

The overlap between the two sets of phenomena is too great to attribute to accident. The particle clitics are surely a subset of the class [41] adverbs. The two groups are alike on semantic, syntactic, and prosodic parameters; the particle clitics further exhibit a special phonological property, liaison, cf. below in section 4.2.2.

Having set forth the claim that the particle clitics are syntactic words, I now show how they fit into the classification of the class [41] adverbs discussed in section 3.1. Interrogative clitic *-kõ* is semantically a conjunction. This is seen most clearly in its 'whether, if' use, as in (19).

(19) En tiedä on-ko isä tulossa.
not know is-Q father coming
'I don't know whether/if father is coming.'

Emphatic clitic *-pA* can be paraphrased by *juuri* 'just', *vain* 'only', and *nimenomaan* 'namely, especially' -- all three are class [45] adverbs. Clitic *-hAn* operates as a connective insofar as it has textual functions (cf. chapter 2.1 and references therein), so it, too, is subset [45]. And clitic *-kin/-kAAn* can be paraphrased by another member of class [45], *myös* (Östman 1977).

Second position clitics like *-hAn, -pA, -kO* appear in the same sentential slot as adverb subclass [41.6] -- those adverbs (*kai,muka*) taking lexical feature [+B]. Particle clitic *-kin/-kAAn* appears in the same sentential slot as its synonym *myös*, that is after the finite verb (a slot described by lexical feature [+D]).

Any separation of particle clitics and adverbs of class [41] will fail to capture the obvious semantic, syntactic, and prosodic overlap between the two. In section 4.1, I argued that clitics are a subclass of dependent words. Here we can say that the particle clitics are a special instance of class [41] leaners, dependent words (not unlike semi-clitic *sitä*) with one further idiosyncratic property: phonological attachment to a host word, that is, liaison.

Liaison is the only factor differentiating the two groups, particle clitics and class [41] leaners. In all other properties, the particle clitics can be considered syntactic words (chapter 2), members of adverb class [41] (chapter 3). The particle clitics, then, are (phonologically) bound words.

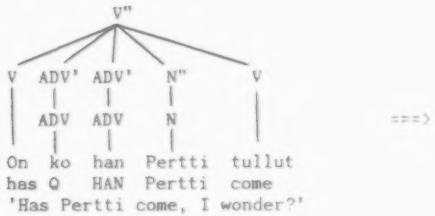
4.2.1. Liaison.

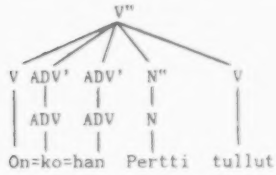
The phonological difference between class [41] leaners and the particle clitics can be attributed to the application of a (post-lexical) rule of liaison (Klavans 1985), which is responsible for the phonological subordination of a syntactic word, such as *-hAn, -pA, -kO, -kin, -kAAn.*

What I have in mind for liaison as a formal operation is merely phonological concatenation. It is not similar to affixation insofar that liaison is not processual -- it cannot, for example, be a stem change, reduplication, or infixal (cf. section 5.4). Liaison subordinates one syntactic word to another. In terms of boundary theory, liaison changes a word boundary into a clitic (or affix) boundary. One previous mention of a liaison rule not given in boundary theory is my "readjustment rule" for the clitic

postposition class in Estonian, whereby a syntactic postposition becomes phonologically part of the preceding noun (Nevis 1982):

In Finnish, liaison concatenates all particle clitics simultaneously:





I take the tack here that one must specify for any liaison operation, the direction of attachment (enclitic or proclitic), as well as the syntactic conditions, of which there are none for the Finnish particle clitics. Other instances of the operation of liaison indicate a need to refer to syntactic parameters like **head of constituent**, **margin of constituent**, and **c-command** (see Kaisse 1985:ch. 7).

The words which undergo liaison are marked in the lexicon by a feature [+liaison], because clitichood cannot be predicted, cf. 4.1 above. Syntax is blind to this feature (as it is to other phonological features), positioning such words as it would other members of [41.6] and [41.8] adverbs. In any event, the direction of liaison has to be specified separately from the syntactic distribution of the particle clitics. (In Finnish all attachment is post-host, as there are no prefixes or proclitics -- see Appendix II).

The lexical feature [+liaison] may have some redundant properties. In both English and Finnish, [+liaison] items are monosyllabic: *-hAn, -pA, -kO, -kin,* and *-kAAn* are all monosyllabic, while the closely related semi-clitic *sitä* is not.

4.2.2. On the Notion 'Clitic'.

There is in the literature on the subject, a serious ambiguity in the term 'cliticization'. Some analysts treat cliticization as a syntactic operation, others as a phonetic operation (cf. Lehiste and Selkirk above). Klavans (1982a:ch. 3) summarizes this problem. She notes a variety of approaches, from clitic (feature) copying, clitic migration, and clitic placement (all syntactic operations), to clitic adjunction or reduction (phonological operations).

In this dissertation, however, I solve this problem, by using the operation **liaison**, which is a phonological operation insofar as it has the effect of creating phonological words. There is no special syntactic operation needed here, because the syntactic positioning of the particle clitics is achieved by independently motivated, more general syntactic operations (within the ID/LP format). This is to say, the syntactic properties of the particle clitics "fall out" from the syntactic properties of class [41] adverbs.

In the same manner I want to avoid the problematic term 'clitic'. First, as I have pointed out in section 4.1, the term has been used to refer to merely unstressed words in a language (i.e. what I have called semi-clitics). And second, clitics do not form a unitary phenomenon (contra Klavans 1982a, 1985). There are two kinds of clitics to be distinguished here: the bound word (appropriate for Finnish particle clitics) and the phrasal affix (appropriate for certain other clitic-like phenomena).

4.2.3. Bound Words: A Problem.

I have been arguing so far that the Finnish particle clitics, and by extension, all second position clitics, are syntactic words. They are located in sentence-second position through the rules of syntax. I have however ignored some of the special properties of second position clitics. In particular, there is a discrepancy among languages as to the actual determination of the notion **sentential second position**. Some languages require that this slot be after an initial constituent; others allow this to be after to the first word in a sentence. And a few languages refer to both determinations for second position.

To a certain extent, Finnish permits reference to both initial constituent and first word, as I noted in 2.1.3. In (20) the clitic does not occur after the noun phrase constituent as expected, but

after the first word *millä*. In (21) the clitic attaches to neither the first word nor the end of the whole constituent, but rather to the head of the constituent (a fact that has previously gone unnoticed).

(20) Millä-hän sivulla se on.
what-HAN page it is
'On what page is it, I wonder.'

(21) kyseessä oleva teos-han, mikä julkaistiin viime vuonna ...
question being work-HAN that one published last year
'The work under discussion, which was published last year...'

The first, (20), resembles topicalization, as discussed in 3.2.7, and I assume that such is the case. The second, (21), however, requires a special rule of syntax. I suggest that a relative clause can be liberated from its NP, so that the relative clause becomes a daughter of the S node. This has the effect of creating the illusion that a clitic is inserted into a constituent -- but what is actually going on here is the location of a second position word after the initial constituent (in (21), *kyseessä oleva teos*) before another daughter of S (in (21), namely *mikä julkaistiin viime vuonna*).

I will not formalize the liberation rule needed for this, but I do point out that the analysis I present here is not unlike English Extraposition from NP, as exemplified in (22).

(22) I want to look the phone number up that I forgot to write down yesterday.

In (22) the relative clause *that I forgot to write down yesterday* happens to be separated (especially in colloquial English) from its head *phone number*.

4.3. On the Phrasal Affix.

Thus far I have focussed on the bound word type of clitic. The Finnish particle clitics are syntactic words which happen to be phonologically attached to some other word. There is nonetheless another sort of clitic in the language, which cannot easily be considered a bound word. This is a morpheme having affixal properties, but with phrasal attachment[6] rather than lexical attachment. I call them **phrasal affixes**.

Phrasal affixes are quite similar to regular inflectional affixes. There are, however, at least two reasons to keep the two separate. First, phrasal affixes attach to one daughter of a constituent, either the rightmost or leftmost daughter (or sometimes

the head). In languages having agreement, phrases will not show agreement for phrasal affixes (whereas they will for inflectional affixes).

Second, even in languages not having agreement, the phrasal affix will lie more remote from the stem than will inflectional affixes.

These two statements hold for both the English possessive *'s* and the Finnish (so-called) possessive suffixes. I consider these two instances of phrasal affixation in some detail in the following.

First, let us consider the English possessive suffix *'s*. The English possessive is clearly a phrasal affix insofar as it does not attach to the head of its NP, but to the end of the NP (Zwicky 1977:7, Janda 1980):

(22) The girl's car

(23) The girl who lives down the street's car

(24) *The girl's who lives down the street

(25) *The girl's car who lives down the street

(26) The woman I interviewed's arguments

In English the possessive morph lies outside inflections, including both nominal and verbal inflections. The external morphotactic positioning of *'s* can be seen in *children's*, where it lies outside the plural mark, and in *interviewed's*, where it follows the past tense morph *ed*.

The Finnish possessive suffixes are affixal in all ways but two: they do not permit the application of the morphophonemic rule of Consonant Gradation (though some of them would be expected to), and they do not undergo adjective-noun agreement (as do other nominal inflectional affixes, for example the case endings):

(27) minun kirja+ni 'my book'
my book +1SG

(28) minun iso kirja+ni 'my big book'

(29) minun iso-ssa kirja-ssa+ni 'in my big book'
*minun iso-ssa+ni kirja-ssa+ni

In Finnish the possessive suffixes do not attach to the margin of the constituent, but to the head noun, and they attach outside all other affixes -- case and number morphs, as above.

There is no reason to lump bound words and phrasal affixes together into one class. The two are qualitatively different. Bound words are treated syntactically as words, which undergo a phonological operation of liaison. Phrasal affixes are clusters of features that must be positioned by syntactic operations (e.g. 'percolation'). Morphotactically, bound words always lie more distant from the host than do phrasal affixes:

BOUND WORD - PHRASAL PREFIX -- INFL. PREFIXES -- DER. PREFIXES - STEM

STEM - DER. SUFFIXES - INFL. SUFFIXES - PHRASAL SUFFIX - BOUND WORD

These facts actually fall out from several pieces of my framework. First, since phrasal affixes are attached already in the syntactic component, but bound words do not get attached to their hosts until after that component (i.e. in a separate component for 'cliticization', cf. chapter 7), the later attachment of bound words predicts greater distance from the host:

after SYNTAX: [word + phrasal affix] [word]

after LIAISON: [[word + phrasal affix] + word]

(and the mirror image for proclitic words and phrasal prefixes.)

Second, this model predicts for the phrasal affix a greater potential to interact phonologically and morphologically with the host than for the bound word. Such is indeed the case (at least) for Finnish (Nevis 1984a). The particle clitics (as bound words) undergo Vowel Harmony with the host, but the Possessive Suffixes (phrasal affixes) undergo other internal sandhi rules as well (including stem formation and allomorphy rules).

The other problem that remains to be solved is the actual syntactic mechanism responsible for producing phrasal affixes (as opposed to regular inflectional affixes).

4.3.1. Phrasal Affixes: A problem.

I envisage the phrasal affix as an inflectional affix with special properties -- frequently appearing are lack of concord and attachment to a non-head. Given the existence of the phrasal affix that attaches at the margins of constituents, we have to account for one universal property that distinguishes the phrasal affix type of clitic from the regular inflectional affix: the external morphotactic attachment of the phrasal affix to the host.[7] Phrasal affixes always lie more distant from the stem than do true inflectional affixes. In the case of the two possessives (of English and Finnish), the phrasal affixes always follow other suffixes in the language. For example, Finnish *kirja-ssa+ni*,

above, has the first person singular possessive suffix after the inessive *-ssa* (with **kirja-ni-ssa* ungrammatical). In the following phrase, English *interviewed's* has *'s* after the past tense suffix (with, of course, the reverse order ungrammatical: **interview-'s-ed*):

(30) The woman I interviewed's arguments

The problem is this: how to capture in the grammar the external attachment of phrasal affixes? For the bound word type of clitic external attachment falls out naturally. Phonological attachment, or liaison, follows all of syntax. Therefore attachment of a word to an inflected word will produce the desired effects:

SYNTAX	[word] [prefix-stem]		[stem-suffix] [word]
LIAISON	word+[prefix-stem]	OR	[stem-suffix]+word

If, however, the phrasal affix type of clitic is to be classed with other inflectional affixes (undergoing, among others, percolation rules), then nothing predicts the ordering of phrasal and inflectional affixes.

4.3.1.1. Solution One.

One approach to this problem would be to place the various clitic parameters -- here, the percolation of a phrasal affix (in Klavans' (1985) terms, the specification of dominance and precedence), as well as its phonological attachment (liaison) -- into a separate cliticization component which follows all of syntax. This is apparently the tack of Zwicky and Pullum (1983), in spite of the fact that placement of phrasal affixes resembles very closely that of inflectional affixes. Although Zwicky and Pullum point out that ...

> Except for the potential distinction between the syntactic locus and the phonological host, such an operation is formally like a rule of agreement -- or more generally, a rule distributing marks of inflectional categories. It follows that special cliticization and inflection can look much alike. (Zwicky and Pullum 1983:511)

Zwicky (1982b) separates cliticization and syntax, positing the former as a distinct component that follows the latter in the syntax-phonology interface. And Kaisse (1985) provides further support. See chapters 5 and 7 for more discussion of this point.

Complicating the clisis component by the addition of inflection-like rules is, all things being equal, simply a way to achieve the external attachment of the clitic type in question.

This approach has no advantage over one in which it is straightforwardly stipulated as a separate statement that phrasal affixes lie outside inflectional affixes. Both solutions I view as equally unmotivated, ad hoc complications of the grammar. Furthermore, both types of cliticization, phrasal affixation and liaison, cannot be in the clisis component, since we will have no way to distinguish the two. And there will be no predictions made as to which type should lie farthest from the host.

4.3.1.2. Solution two.

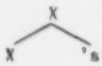
A less unmotivated approach is one that follows from the principles already needed for immediate dominance and linear precedence in GPSG. The marginal location of affixes will be determined in the same fashion as the location of affixes on heads. Feature percolation will proceed as in regular inflectional morphology, through feature-passing conventions. But, for phrasal affixes the features are transmitted from the mother category to exactly one of the daughter categories (which may or may not be the head daughter). We know from linear precedence statements that syntax must have recourse to such notions as 'first' and 'last' in a constituent, cf. chapter 3. The percolation of a feature can thus be linked to one of the following: [HEAD], [FIRST], [LAST].

So far the location and percolation of phrasal affixes parallels closely the percolation of inflectional lexical affixes (except in the linking of the feature(s) to [HEAD], [FIRST], or [LAST]). The difference between the two types of affix lies in where the feature complex is realized as morphological material/processes. Inflectional affixes get realized in the morphological component whereas phrasal affixes are realized in the syntactic component. Such a distinction may appear ad hoc, but actually the distinction captures another generalization to be found in the grammar -- phrasal affixes pattern more with other grammatical function words (i.e. adpositions, adverbs, conjunctions, etc.), which are likewise realized in syntax.

For the syntactic realization of features, a precedent has been set by Gazdar's (1980) treatment of English dative *to* (as required for subcategorization by verbs) as a feature, [to]. He also treats the conjunctions *and*, *or*, etc. as features. In both groups of examples, the feature complexes get realized syntactically as (grammatical) function words rather than morphologically as inflectional affixes/process morphology.

English possessive *'s* will now be linked to the feature [LAST], and thus must occur on some lower branch of the NP which has inherited the feature. It is not simply a feature on that daughter constituent, otherwise we cannot distinguish among lexical inflectional features and phrasal features. In the case of the phrasal affix, an additional branching is required.

This branching must be part of ID, so I propose that English 's, for example, is represented in the following structure:



Structures like this one containing phrasal affixes,require a special ID rule which refers to margin features.

<10, X ---> X, [MARGIN:a]>
[MARGIN:a]

Here [MARGIN] represents the two margin features [FIRST] and [LAST]. [a] stands for a particular linked morphosyntactic property, such as the English possessive. This rule has the effect of creating the equivalent of a grammatical function word, yet at the same time chomsky-adjoining it to its lexical host.

Additional LP statements are required to ensure that the newly created margin affix precedes its host if the affix is [FIRST], but follows if the affix is [LAST].

X < [LAST:a]

[FIRST:a] < X

The attachment of an affix like a Finnish possessive suffix, which is linked to [HEAD], cannot be distinguished from other head features, ones that show agreement, such as case and number. So, case, number and possessive features are all head features, but possessive further requires branching. Branching thus ensures the outer location of phrasal, as opposed to lexical, affixes.

This scheme has several advantages over the two presented above in 4.3.1.1 -- (1) the non-interaction of phrasal and inflectional affixation falls out of principles independently motivated in GPSG for the distribution and percolation of features; (2) the similarities between inflectional morphology and phrasal affixation are captured in the notions **dominance** and **percolation** (with certain differences, of course); (3) the difference in morphotactic location of the two classes of morphemes is captured in the additional branching needed for phrasal affixes, and parallels between phrasal affixes and grammatical function words come out here; and (4) the distinction between phrasal affixes and bound words is not blurred. In fact, we can see now that there are very few shared properties in the two types of clitics (bound word vs. phrasal affix), with two exceptions -- location at word margins and (at least some) similarity to full words.

One final advantage for my proposal is the prediction that inflectional affixes will be numerically predominate in the languages of the world. Phrasal affixes will tend to be rarer because they will require some additional machinery in the syntax.

4.4. Conclusion.

In this chapter I have argued that the Finnish particle clitics are phonologically bound words, which are a proper subset of semi-clitics (or 'leaners' in general). Bound words and leaners share semantic, syntactic, and prosodic properties, but differ in phonological behavior: bound words are leaners displaying additional phonological interaction with some host. Phonological interaction is the result of the application of a rule of liaison, or phonological readjustment.

Bound words are to be kept distinct from another type of clitic, namely 'phrasal affixes'. The two are distinguishable on syntactic. morphotactic, morphological, and phonological grounds. The bound word is a syntactic word (lexeme); morphotactically most distant from the host; morphologically discrete material; and phonologically barely interactive with the rest of the word. The phrasal affix, by contrast, is a syntactic affix; morphotactically less distant from the host than the bound word (but more than other affixes); morphologically heterogeneous; and phonologically highly interactive with the host.

Qualitatively, too, the two sets of clitic phenomena are distinct. They each call for different 'cliticization' operations: bound words undergo liaison; phrasal affixes are attached already in the syntax as feature complexes (and require a special branching from the host).

Endnotes.

1. Others include Fudge (1969), who distinguishes the grammatical hierarchy from the phonological hierarchy; Pulgram (1970) who makes a similar distinction (he calls the phonological word nexus); Selkirk (1984:ch. 6); and the autosegmental phonologists.

2. Actually, he considers the inclusion of such stressless words a syntactic-semantic unit, rather than just a phonological phrase -- his term is *solu* 'cell'. The main difference between the *puhetahti* and the *solu* is whether a phonological phrase begins with a main stress or not. E.g. *ja-kóko Suomen kansalle* 'and to the whole Finnish people' begins with stressless *ja* so is not a "puhetahti", but is a "solu". In other words, *solu* = *puhetahti* + satellites.

3. By the term 'precision of timing', Niemi (1984) means the inverse of variable timing: greater variation underlying some arithmetic mean of some unit in speech production corresponds to difficulty in controlling the timing of that unit; and conversely, less deviation corresponds to greater control in the timing (Niemi 1984:61ff). Niemi defines his index for precision as "the simple arithmetic mean divided by standard deviation" (p.62). It is relative to a base value.

4. A partial answer to Sadeniemi's queries can be found in Leino's (1982) book on Finnish metrics. Leino offers a number of prominence rules for the language. He attempts to predict, for example, when phonological pauses are possible between constituents and when main stress is permitted on a grammatical function word. Leino labels the negative verb, pronouns, conjunctions, and certain adverbs as 'grammatical morphemes'. He further observes that when two normally stressless grammatical morphemes occur side by side and when main stress should be assigned to one of them by other principles, there are three possibilities: neither takes main stress (a); the first takes it (b); or the second takes it (c).

a. Nyt taas sa̋veleet sůlokkáiset
   now again melodies sweet
   'Now again the sweet melodies'

b. Tuo kátse, -- nýt mä̆ tiedä̆n, ...
   that glance now I know
   'That glance, -- now I know ...'

c. Ah, jós rikas, nůorikin ó̆isin
   oh if rich young-too would be-1SG
   'Oh, if rich, I would be young, too'

Leino's prosodic rules are rather sketchy, and are mainly intended to be applied to Finnish poetry. Consequently I do not explore them here.

5. In yet other dialects, conjunction clitic *-jA* is not a conjunction, but serves as a replacement for clitic *-kin*, i.e. it has the meaning 'also, too', as in the sentence *Minä tulen-ja* 'I am coming, too'. Examples can be found in a variety of dialects, whose descriptions are located in the Dialect Archives, **murrearkisto**, at the University of Helsinki.

6. I should also cite here Ravila's (1941:30-34) mention of the German term **suffixlockerheit** used by Ernst Lewy in reference to the dative *-a* and instrumental *-at* of Ostyak.

a. kur-en uč-en-*a* (*kur-en-a uč-en-*a*) 'zu deinen Füssen, deinem Gewand'
b. täbettäiŋen, jäŋǥeltäiŋen sar-jiŋk, mag-jiŋ*at* (*... sar-jiŋ*at*, mag-jiŋ*at*) 'sie wurden genährt, getränkt mit Bier, mit Met'

These examples show that the dative *-a* and instrumental *-at* can have apparent phrasal attachment.

Suffixlockerheit is not quite the same as phrasal affixhood, but may be used as a good indication of it.

7. The external attachment of clitics is seen in the overwhelming predominance of enclitics and proclitics, but the relative scarcity of endoclitics. See chapter 5.4 for discussion.

CHAPTER V

ON CLITICS AND CLITICIZATION

5.0. Introduction.

Previous work on the Finnish particle clitics is suggestive of a treatment of them as syntactic words. Genetz, L. Hakulinen, and Setälä, among others,[1] explicitly consider the particle clitics to be members of adverb (and conjunction) word classes. Genetz (1892) said that the cliticized particles are semantically adverbs and conjunctions, and are positioned after the most stressed word of the sentence, as a modifier of it.

> The cliticized particles are, according to their meaning, words on the border between adverbs and conjunctions, which one is accustomed to using after the most stressed word (usually the first) of the sentence, as a modifier of it, and which have, through that, lost their stress and adapted themselves to vowel harmony with the preceding word. (Genetz 1892:93)

Genetz placed the particle clitics in the same section of his book as the adverbs, so he probably considered them to belong to the same word class.

L. Hakulinen (1979) treats the particle clitics as full words. He calls the clitic *-hAn* as adverb in Standard Finnish, but also a copulative conjunction in the colloquial languages (p. 75). The question clitic *-kO* he equates with the independent (dialectal) conjunction *ko* (as in *se ko tahtoo = se kuka tahtoo = se joka tahtoo* 'he who wants', p. 88). In general Hakulinen describes the clitic particles as adverbs (unless otherwise specified, see p. 235). The adverb clitic particles, then, are *-s, -hAn, -kA* (in some uses), *-kAAn, -kin,* and *-pA*. Clitics *-kO* and *-kA* (again, in some uses only) are labeled conjunctions. Hakulinen notes that the close relative of *-kO*, independent *ko* in the dialects, does not participate in

vowel harmony: *hänko* is used in place of standard *hänkö*. This makes *ko* a semi-clitic (like *sitä*) rather than a full clitic. This is to say, dialectal *ko* is a strictly positioned, unstressed conjunction lacking the phonological interaction with the host that signals a true clitic.

Setälä likewise treated the particle clitics *-kO* and *-hAn* as adverbs (1930:120).

Penttilä (1963:115ff) has perhaps the most extensive and most theoretically oriented treatment of the particle clitics. He distinguishes among **sane** 'lexeme', **sananmuoto** 'word-form', and **sana** 'word' (as well as **tyvi** 'stem'). The **sane**, or lexeme, is the abstract, fundamental unit of the lexicon. When lexemes are realized concretely in phonological form, we are dealing with the **word-forms** of the lexemes. The term **word** is probably meant to be a conglomeration of lexemes and morphemes (à la Matthews 1974:26). Penttilä considers the particle clitics to be lexemes separate from the host lexeme. Thus, the word-form *tulenpas* 'I am coming' contains three lexemes -- *tulen* (composed of lexeme *tule-* 'come' and morpheme *-n* 1SG), *-pa*, and *-s*. He also considers instances in which specialized, lexicalized clitics are to be viewed as inflectional or even derivational morphology (cf. his remarks on *ku-ka*, *mi-kä*, and *kuten-kin*, p. 119).

Penttilä's inventory of particle clitics is much broader than mine. He includes not only *-hAn*, *-kin/-kAAn*, *-kO*, and *-pA*, but also *-kA*, *s* and dialectal clitics such as *-mA* and *-stA*. He further lists clitic combinations. Lexeme compounds include *jopa ... -kin 'even ... too'*, *jos ... -kin 'if ... also'*, *niin ... kuin ... -kin ' both ... and ... also'*, *niin ... jos ... -kin '(both) ... if ... too'*, *niin hyvin ... kuin ... -kin 'both ... as well as ...'*, *vai ... -kO* (also *-kO ... vai ...*)'or?' (p. 556-559).

Penttilä remarks (p. 555) that Finnish has a number of 'questioning sentential modifiers' of which some are interrogative conjunctions. The particle clitic *-kO*, for example, is listed alongside several full words.

Of the scholars in the younger generation, most are interested in the function of the particle clitics in discourse. F. Karttunen examines the particle clitics' interactions with constituent order, ascribing to *-hAn* an illocutionary force. A. Hakulinen, J. O. Östman, M. Vilkuna, and M. Vilppula take similar lines. They argue that the particle clitics straddle the boundary between syntax and pragmatics. Generally they relate pragmatic clitics like *-hAn*, *-pA*, *-kin/-kAAn* to other pragmatic markers (full words) like *nimenomaan* and *myös*, so their work is entirely compatible with my own.

F. Karlsson, in various publications, looks at the particle clitics from the point of view of a morphologist. He shows on a number of criteria that the particle clitics are not in the least affixal, cf. chapter II. He does not demonstrate that the particle clitics are bound words, but his results are definitely compatible with such an approach.

The above discussion reveals that the insights of native Finnish linguists favor the bound word approach. Many of the earlier investigators explicitly included particle clitics alongside presentations of regular words. And the work of more recent investigators is consistent with my bound word approach.

For the rest of this chapter I focus on the development of a theory of cliticization. Earlier clitic taxonomies are presented in 5.1. Three approaches to the status of clitics are discussed in 5.2. Some syntactic and phonological parameters of cliticization are explored in sections 5.3 and 5.4. Finally, a few diachronic studies of second position cliticization are mention in section 5.5.

5.1. Clitic Taxonomies.

5.1.1. Introduction.

A binary division of labor for all clitics was posited in chapter IV. Clitics will be either **bound words** or **phrasal affixes**. This is a rather novel taxonomy, though not entirely without precedent. There are some hints in the literature to the effect that clitics may be considered one or the other, but these remain inexplicit and language-particular. Nida (1946:104-6), for example, cites English genitival *-'s* as an affix that occurs with phrases, but he does not actually consider this a clitic (nor does he recognize any bound word phenomena *per se* -- instead he lists separately, "bound alternates of free forms" and "clitics which are not relatable to free alternates" (p. 106).

In the following three sections, I take a look at several clitic taxonomies that have been proposed in recent years.

5.1.2. Zwicky 1977.

Zwicky (1977) presented one of the first works pulling together a unified account of clitics from a variety of languages. He posited a tertiary division among clitics: **the simple clitic, the special clitic, and the bound word.**

The simple clitic is defined as a morpheme that has a free alternate. There is a transparent phonological and syntactic relationship between the two alternants. English auxiliary verbs constitute examples of simple clitics. English *is, has, have* etc. have single consonant, clitic allomorphs /z/, /z/, /v/ beside

full /əz/, /hæz/, /hæv/, respectively. The syntactic constraints on 'contraction' (i.e., cliticization) are well-known (cf. summaries in Kaisse 1983, 1985; Bissantz 1983/1985), as is the fact that phonological attachment is to the left, as demonstrated by voicing assimilation.

(1) John[z] taking -- *John[s] taking (i.e., *not* John staking)

Pronouns in many languages have full and simple clitic allomorphs under Zwicky's (1977) analysis.

Special clitics, according to Zwicky (1977), also have free alternants, but are characterized by idiosyncratic syntax and/or irregularities in phonological shape (in comparison to the free form). An example of this is the French pronominal clitic *le* 'him'. It differs from its free alternant *lui* in both parameters -- special phonology ([l(ə)] is not relatable to [lɥi] through any productive synchronic phonological rule) and special syntax (*le* has a specific requirement to be adjacent to the verb; *lui* is not always permitted in those slots -- cf. (2-4) below).

(2) Je le vois -- *Je vois lui -- *Je lui vois
I him see
'I see him'

(3) Donne-le-moi -- *Donne-lui-moi
give-him-me
'Give it/him to me'

Finally, Zwicky places all clitics without any known free-form alternant into the **bound word** category. Latin -que, -ne and all the Finnish particle clitics are examples of bound words, because there are no free-form alternants of these morphemes.

(4) Latin a. navigia mittantur, deis-que iuvantibus [illegible]enient
ships should be sent gods-and helping will arrive
'Ships should be sent, and with the gods' help, they will arrive.' (Jansson 1976:238)

b. etiam-ne nobis expedit?
really-Q for us useful
'Is it really useful to us?

Klavans (1982a) has criticized Zwicky's (1977) tertiary division on the grounds that the special clitic/bound word distinction loses generalizations. She notes that there is in Zwicky's treatment a strong correlation between pronominal clitics (which very often have parallel 'strong/weak' -- i.e. full/clitic -- allomorphs) and special clitic status, as well as between particle

clitics (usually adverbial or complementizer clitics that lack 'strong/weak' allomorph pairs) and bound word status.

Klavans also observes that the phenomenon of second position enclisis is not captured in any straightforward manner under Zwicky's (1977) approach. Some second position clitics are treated as special clitics, others as bound words. Ngiyambaa, for example, has both pronominal and particle second position enclitics. Presumably the former would be labelled 'special clitics', the latter 'bound words', if we follow Zwicky's 1977 taxonomy. The shared syntactic behavior does not fall out of this early approach of Zwicky's.

5.1.3. Pullum and Zwicky 1983.

Klavans' (1982a) criticisms of Zwicky 1977 led to Zwicky and Pullum's (1983) revision in favor of a binary clitic division. Simple clitics have the same status as before, but now the term special clitic is used to cover both of Zwicky's (1977) categories, special clitic and bound word.

Simple clitics, then, have exactly the same privilege of occurrence as that of associated full forms. The only mechanism required for these clitics is a readjustment rule (i.e. liaison):

> the formal device which creates phonological words containing a simple clitic is a readjustment rule operating on the surface structure. (Zwicky and Pullum 1983:510)

This is precisely the proposal I offer in section 4.2, as my rule of liaison.

But Zwicky and Pullum lump together all other clitics under the special clitic type:

> All other clitics are special clitics in our terminology (Zwicky 1977): either no corresponding full forms exist, as in Latin conjunctive particle *-que*, the Tagalog clitic particles, and the English possessive *'s*; or else the clitics do not have the same distribution as the corresponding full forms, as in the pronominal clitics of many Romance and Slavic languages and of Modern Greek. (Zwicky and Pullum 1983:511)

Thus they place bound words like Finnish particle clitics and Latin *-que* (as well as the Tagalog [semi-]clitics) into the same category as phrasal affixes like English *-'s*. This approach

will miss generalizations about the bound word/phrasal affix distinction or will predict that bound words and phrasal affixes should pattern together. But as I note in section 4.3, there is a qualitative difference between the two phenomena.

Zwicky and Pullum (1983) view special clitics as analogous to affixes -- the two groups of morphemes are to be represented as features and are to undergo rules of 'percolation' in order to be positioned syntactically (cf. 5.2.2 below). But this feature approach blurs the distinction between bound words and phrasal affixes. Under their view, a clitic such as the English possessive marker, *-'s*, is treated in exactly the same manner as the Estonian clitic, or bound, postpositions (Nevis 1982). The clitic status of the Estonian bound postpositions is motivated, not on their affixal properties, but on the syntactic parallelism with regular postpositions. English *-'s*, however, can be motivated as a phrasal affix (cf. section 4.3) and should then be handled via a syntactic feature.

Another such contrast can be found in a single language, namely Finnish. Finnish has two classes of clitics -- bound words and phrasal affixes. These are both special clitics according to Zwicky and Pullum's (1983) approach: they either have no free allomorph (as is the case with the particle clitics) or else there is no transparent syntactic and phonological relationship between the clitic and the free form (as is the case, for instance, with possessive enclitic *-ni* and its free allomorph *minun*, both meaning 'my'). A special clitic treatment of both the word-like clitics *-hAn, -pA, -kO,* etc. and the affixal possessive clitics *-ni, -si, -nsA,* etc. fails to distinguish between morphemes that call for a feature/affixal approach and those that call for a bound word approach.

Several recent studies of clitics make heavy use of this simple/special clitic taxonomy, including Kaisse (1985) and many analysts working within the government and binding framework (e.g. Saxon 1985).

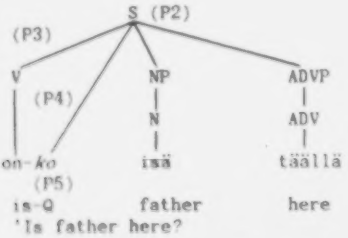
5.1.4. Klavans 1982a.

Klavans adopts a different approach from Zwicky (1977); she explicitly assumes that clitics constitute a unitary phenomenon. According to Klavans, clitics are the same 'thing', though subject to five parameters which take care of the clitics' syntactic and phonological positionings (P1-P5 below). There are no types to be distinguished (i.e. there is no clitic taxonomy *per se*); instead, variation arises as a result of the interaction of the parameters.

P1. Clitic Identity
P2. Domain of Cliticization
P3. Initial/Final Location within that Domain
P4. Before/After the Host Word
P5. Proclitic/Enclitic Attachment to Host Word

Parameter 1, clitic identity, is represented by a lexical feature [±clitic]. Since clitichood cannot be predicted (cf. section 4.1), clitics are marked [+clitic] in the lexicon. Parameter 2 refers to the node that dominates the clitic -- the clitic is positioned syntactically with respect to the daughters of that node. Parameter 3 refers to the margin of the domain -- is the clitic placed before all the other sisters (i.e. left margin), or after the other sisters (i.e. right margin). The fourth parameter determines the locus of clitic attachment with respect to its host word. And parameter 5 dictates the direction of phonological attachment -- the clitic will either be enclitic or proclitic.

To illustrate the use of the five parameters, I have selected Finnish particle clitic *-kO*. Parameter 1 will obviously obligate us to mark this word as [+clitic]. Its domain is that of the sentence (i.e. it has sentential scope), so P2 = the *S* node. *-kO* does not appear final in the sentence; rather it refers to the left margin (my feature [FIRST] in chapter 3), so P3 = initial in Klavans' terms. The clitic *-kO* appears after the first daughter constituent in *S*, so P4 = after. And finally *-kO* is enclitic, not proclitic.



This is the typical combination of parameters for the phenomenon of second position enclisis (also known as **Wackernagel's Law**).

In the IULC edition of her dissertation (in a section of the introduction entitled "How Five parameters Became Three"), Klavans points out that only three, not five, parameters are actually needed to capture the positioning of clitics. The lexical feature [+clitic] is really a part of the lexical entry of any lexical item. Parameter 2 is redundantly stated in parameters 3 and 4: P3 refers to initial/final location within a particular domain, and P4

to location before or after a particular daughter of that domain. The three parameters now are:

P1. Initial/Final
P2. Before/After
P3. Enclitic/Proclitic

The above revised parametric system, leads ultimately to Klavans' (1985) article in which she posits the same three parameters, but with different labels:

P1 = Immediate Dominance
P2 = Linear Precedence
P3 = Liaison

I am in complete agreement with Klavans here, concerning the parameters for 'cliticization'. But I question whether P1 and P2 are truly parameters for clitics. It seems to be no accident th.. Immediate Dominance and Linear Precedence are needed for both syntax and cliticization. The term 'immediate dominance' is taken from GPSG, where it likewise refers to the configurational structures of units -- a domain (or node) has several members (or daughters). And linear precedence refers to the linear ordering of the daughters of that node (or actually, of the sisters under any shared node). The only parameter that is not syntactic in nature is P3, which is the direction of phonological attachment.

In particular, we note that the Finnish particle clitic *-kO* has the following parameters, under Klavans' 1985 treatment: ID = initial in S; LP = after; and P3 = enclisis. Klavans' approach has the unfortunate consequence that the clitic parameters recapitulate the same syntactic ID/LP statements for class [41] sentential adverbs. Klavans thus introduces a large amount of redundancy into the componential organization of the grammar. The syntax and cliticization modules are governed by the same operations: ID and LP. I postpone my discussion of the implications of permitting such redundancy among components until the conclusion of this dissertation (chapter VII).

It should be noted that I am not taking issue with Klavans' parameters *per se*, only with the componential allotment of the parameters. We can cut out repetitive operations between these two modules by allowing clitic ID/LP to fall out of syntactic ID/LP. I proposed in chapter 4 that Klavans' ID and LP be relegated to the syntactic component.

The coupling of Klavans' (1985) parameters with my proposed revisions now reduces the notion 'cliticization' to liaison, or direction of phonological attachment. Under this revision, there are just two types of cliticizations: proclisis and enclisis. All other parameters are determined by the syntax of the language in

question. I have adopted Klavans' lexical feature [±clitic] in section 4.1.1 where I have renamed it [±liaison] to emphasize the fact that the feature involves no syntactic operation.

5.2. Three Views of Clisis.

Analysts' views of the notion 'clitic' can be reduced essentially to three basic approaches toward the status of clitics in grammatical theory. First, there are linguists like Steele (1975) who consider clitics to be linguistic primitives. Then, there is the clitic-qua-affix camp, represented by Klavans (in various publications, cf. 5.2.1 below). Klavans argues that clitics should be seen as affixes with phrasal determination. And finally, there is the bound word approach typical of Finno-Ugricists (section 5.0 above). In the following I examine these three viewpoints, arguing that none of them is sufficient alone to handle the array of data available in the literature. I offer instead a binary approach to clisis: clitics will either be bound words or phrasal affixes, depending on the outcome of Zwicky's (1984a) diagnostic tests.

5.2.1. Clitics As Linguistic Primes.

One view of the status of clitic, that of Steele (1976), dictates that clitics are linguistic primes. (See also Comrie's remark in 5.5.1.) Steele comes to this conclusion after being unable to find any motivation for second position enclisis. She says, "asking the question: Why second position? is like asking the question: Why adjectives?" (Steele 1976:560).

This view does not have much to recommend it. If clitics are viewed as basic units of language, then one must determine how such units fit into linguistic theory. First, the notion must be defined in such a way as to indicate its intermediate status between the word and the affix. This is especially crucial for agglutination facts, according to which diachronically words typically become clitics on their way to affixhood.

Numerous affixes have arisen through the reduction of full words. Most frequently there is not a direct reinterpretation of a full word as an affix. Instead there is an intermediate stage during which the morpheme in question is a clitic:

WORD > CLITIC > AFFIX

The proponents of the linguistic prime approach must elaborate on how such a primitive of language will fit into synchronic and diachronic studies, child language acquisition, borrowing, aphasia, language play, and the like. It is further not clear to me that such a fundamental unit should have two and only two subtypes -- the bound word and the affix. Are these each to be considered separate units of grammar? Or is there some reason to have binary division for this linguistic prime?

I take a very different approach here. A clitic is to be defined in terms of other, independently motivated linguistic primes, namely the word and the affix. Its intermediate status between that of the word and that of the affix can be demonstrated with Zwicky's (1984a) tests. Not just any combination of word and affix properties is available. Rather, clitics exhibit a clustering around two poles -- word properties OR affix properties. Under my view, then, the clitic is either a type of word or a type of affix, and such clusterings lead to my bound word/phrasal affix distinction.

In this dissertation I have taken the tack that the addition of some new basic unit or some new grammatical apparatus is to be avoided. I have attempted to force as much as possible to fall out from machinery already available in the rest of the grammar.

The analysis I have offered relies instead on independently motivated mechanisms, with the exception of the lexical feature [±liaison]. This feature is needed in order to account for the inability to predict clitichood on the basis of other factors (such as stresslessness, semantic vagueness, and the like). So, this feature is one small piece of the grammar (not an entire component, or set of percolation rules) that cannot be reduced to any existing principle. But the feature serves to signal the patient which will undergo phonological liaison (i.e. marks a lexeme which will become bound).

5.2.2. Clitics as Affixes.

Groos (1978) attempts to create an inflectional theory of clitics, based on data which highlight the similarities between Spanish and French pronominal clitics and inflectional affixes. Stump's (1980) detailed account of the French clitics places such data in doubt as examples of true clitics. Stump provides an inflectional account of the French pronominal and adverbial clitics. He argues that the clitics are represented as morphosyntactic inflectional features on the verb, and that the clitics are not derived from corresponding full noun phrases or prepositional phrases.

Zwicky and Pullum (1983) present arguments for English *n't* as an inflectional affix, ending their article with a short discussion of the parallels between affixes and clitics.[2] They point out that the percolation of clitics parallels that of affixes:

> Phonological words containing a special clitic could be regarded in transformational terms as created by a rule that takes features associated with some domain constituent (usually S or NP); transfers them to a locus, a specific node within the domain (e.g. to an initial or final sub-constituent, or to the head of the

> constituent); realizes them as morphological material situated either before or after the locus; and attaches this material phonologically either to the right or to the left (p. 510).

This rather complex scheme, based on Klavans (1982a), no doubt captures the relevant positionings of clitics. But it has some undesirable properties, most notably duplication of function and missed generalizations. I have argued in section 5.1.3 that, for the bound word type of clitic at least, 'percolation' rules of this type will recapitulate what already has to be done in the syntax of the language. Having one set of percolation rules for all clitics further blurs the distinction among the two clitic types -- whether that distinction be special/simple clitic or bound word/phrasal affix.

Finally, I will argue in chapter VII that only liaison is necessary to handle cliticization. Zwicky and Pullum's percolation approach, like Klavans' ID/LP approach, makes the cliticization component too powerful. Allowing syntactic operations in that module mixes phonology and syntax in a way that weakens the theory.

On the other hand, Klavans (1979) notes several times that there exists a distinction between clitics having categorial membership and those having phrasal attachment (p. 62-63), and she specifically cites a distinction in Turkish between enclitic words and enclitic suffixes (p. 98ff). In later work, however, she takes a different stand: she explicitly states that clitics are best viewed as "phrasal affixes" (Klavans 1982a:xvi-xviii; 1983:104; 1985).

The phrasal affix analysis is clearly not suitable for the Finnish particle clitics, as these morphemes are much more word-like than they are affixal. But the concept is not to be thrown out entirely, as it is needed for clitics like English *'s*, as I have shown in chapter V.

5.2.3. Clitics As Words.

Several investigators take the approach that clitics are words. Hale (1972) presents such an account of the Warlpiri auxiliary. Steele and Demers (Steele et al. 1981) provide a similar view of the Luiseño and Lummi auxiliaries. And Klavans' most detailed account of clisis comes from her (1983) article on Ngiyambaa. Ngiyambaa has amazing freedom in its constituent order, but has a severe restriction on the placement of sentential clitics into the second position of the sentence. Klavans offers the following fragment of a phrase structure grammar for the language:

```
S ---> a (ENCL) a*
ENCL ---> PART[icle] PRO[nominal]
```

Klavans' rules probably capture the surface order of the language, but in a base rule such as this, the use of the ENCL[itic] node is unwarranted insofar as the clitic is not on a par with other syntactic elements. A clitic is not a word class, but encompasses other word classes, such as "adverb, pronoun, and "any morpheme class ... save (lexical) verb and (lexical) noun" (Klavans 1979:36). I agree with the position argued by Zwicky (1984a) that all words are to be assigned a word class and that there are no acategoremetic words. Thus PRO[nominal] and PART[icle] may be used as word classes, but ENCL[itic] may not. Nor can it be used as a phrasal node. It is true, however, that sometimes it is difficult for the analyst to assign a word class, especially in the absence of inflectional morphology on the word (as is often the case for clitics).

Another criticism of Klavans' phrase structure rules above follows from an observation made by Zwicky (1982a), who notes that clisis rules never feed or bleed syntactic rules (an observation which leads to his positing a separate post-syntactic cliticization component). Klavans' (1983) phrase structure rule is a glaring instance of a mixing of levels: reference to enclisis is included among the base structure rules of the language.

In an earlier work, Klavans (1979) demonstrates that at least some clitics are to be considered members of "inflectable word classes" such as noun and verb. This is because we can account for the surface phenomenon of endoclisis if we assume that the inflectional affix that appears external to the clitic+host unit is not attached to the host, but to the clitic itself. Thus, in (5), from Ngiyambaa, the suffix *-bula* (a duality marker) is affixed to pronominal clitic *-ndu*, which is itself attached to the host *dhi:rbawa* (with intervening clitic *-nha*):

(5) dhi:rba-wa -nha -ndu -bula ŋiyamba: niya-li
know-GETTING-PRES-2NOM-DUAL Ngiyambaa speak-PURP
"You two are learning how to speak Ngiyambaa.'
(Klavans 1979:71)

Suffix *-bula* never attaches to verbs like *dhi:rbawa*, but always attaches to nouns and pronouns: *bura:y-bula* '(the) two children' and *nindu-bula* 'you two'. *-ndu* is the clitic form of *nindu*, and *-ndu-bula* is then the clitic form of *nindu-bula*.

Klavans concludes her interesting paper with a statement she will later contradict:

> In this paper I have argued that some clitics are underlyingly words on the basis of their class membership and their inflectability, and that the analysis of clitics as morphological features is untenable. I have

> proposed that clitic words are always attached externally to hosts and joined by Cliticization Rules, that is, rules which create phonological words from grammatical words. (Klavans 1979:77)

As my discussion of the Finnish particle clitics demonstrates, I am in wholehearted agreement with this statement. But not all clitics are to be considered bound words (cf. sections 5.2.2 and 4.3).

5.3. On the Syntax of Clitics.

Most investigations of clisis reveal that clitics tend to attach to the margins of constituents (Zwicky 1977, Klavans 1982a, Kaisse 1985). But I argue below that reference to the head of the constituent is also an available option. In this section I address and criticize Kaisse's very strong contention that sentential clitics must occur in clausal second position.

5.3.1. Heads and Margins.

Several previous studies emphasize the fact that clitics are located either at the edges of a constituent or attached to the head of the constituent. Zwicky (1977) and Klavans (1982a), especially, argue this point. Zwicky notes that "clitics whose source is within a particular constituent (an NP or S) will move either to one of the margins of that constituent or to the head of that constituent (the N or V)" (p. 18).

Klavans echoes that view, and notes that there is a great deal of overlap between Zwicky's discussion of the location of clitics at the margins or head of a constituent, and Baltin's (1978) view that 'movement' rules in Chomsky's 'move alpha' theory of syntax refer to only to margins (and head) of a constituent.

Kaisse (1982, 1985) formalizes Zwicky's statement, and uses it as a more general principle for her second position rule governing the syntactic location of sentential clitics (see below in section 5.3.2):

> A clitic originating as an immediate daughter X^n may [be] adjoined only to the leftmost or rightmost node of X^n. (Kaisse 1982:5)

In a footnote she appends the possibility of the head of the constituent serving as the host:

> ... or to the leftmost or rightmost node of X. (p. 13)

Kaisse's notions of constituent and head are purely syntactic -- they do not hold any special non-syntactic status. That is, the two notions do not behave any differently in the cliticization component than in the syntactic component.

Klavans (1982a:preface) takes a different approach to the syntactic location of clitics. She wants to argue that clitics are a type of phrasal affix, so she assigns subcategorization frames to them, just as Lieber assigns such frames to affixes. But as I have argued, many of the clitics are not affixal at all; they are instead bound words.

5.3.2 On Wackernagel's Law.

Kaisse (1982, 1985) examines second position, or Wackernagel-type, clitics within the (pretheoretic) framework of Klavans (1982a), and attempts to impose restrictions on it. Using a corpus of only six languages, she observes that there is a significant overlap between sentential clitics and second position clitics. Sentential clitics have a very strong attraction for this sentential slot. In Klavans' (1982a) terms, the following properties cooccur: sentence initial, after the host, and enclitic onto the host.

Kaisse defines sentential clitics as the immediate daughters of S or S'. She then formalizes her very strong claim:

> All those languages with S' clitics place those clitics in second position, after the first stressed constituent (or word) of that clause, regardless of the category of that constituent (or word). (Kaisse 1982:4, footnote deleted here).

Unfortunately two problems crop up for Kaisse's claim that all languages with sentential clitics place them in second position. First, although she intially cites six languages which obey the rule above, she mentions almost as many languages in which the sentential clitics do not appear in second position: Kenyang, Ngancara, Welsh, and Chrau.[3] In the discussion that follows, I present several other counterexamples to Kaisse's claim from the Finno-Ugric language family.

Kaisse also fails to capture the notion 'sentential clitic'. She offers a (universal) structural definition -- whatever is immediately dominated by the S or S' node. Kaisse can account for adverbial sentential clitics with this definition, as well as conjunctions and complementizers. The definition also covers subject pronouns, but leaves out other pronouns (e.g., direct and indirect object pronouns). Other problematic sentential clitics include auxiliary verbs (which in English are only controversially analysed as immediate daughters of S) and tense markers. In my syntactic account of Finnish all the above are daughters of S (or S'), and such a treatment could be extended easily to other free word order languages.

But in Kaisse's treatment, many pronouns, such as objects and indirect object pronouns, as well as genitive pronominal clitic attributes of a predicative NP, are daughters of the VP node, not of the S node. In order to handle the second position placement of these pronouns, she has to append a rider -- the second position clitic group can be extended by analogy. The positioning of the subject clitics is generalized to include other pronominal clitics.

Kaisse's law is empirically false. There is no exceptionless principle that forces clitics into sentential second position. This can be seen in the following data from Finnish, Veps, and Permic-Ugric. In these languages, the finite verb holds a very great attraction for sentential clitics.

I have collected a corpus[4] of nearly fifty languages having second position clitics, all of them exhibit a great amount of free constituent order. Some of these languages are, in Hale's (1981) view, W-star, or non-configurational languages, others are not. They all presumably have flat syntactic structures like those proposed for Finnish in chapter III. Sentential adverbs and conjunctions, auxiliary and finite verbs, and various sorts of pronouns would all be daughters of S (or S'). The only remaining problematic instance is the genitive pronominal clitic attribute of a predicative NP (cf. (6) from Pashto, where *de* is the genitival attribute of *wror*, and is in the middle of a cluster of second position [semi-]clitics). I cannot account for this loose thread at the present time, but I note that there is a strong correlation here between ergative pronouns and these genitive attributes. This may simply indicate a flatter structure for noun phrases in these languages. (See also Zwicky (1985c) on similar flat structures for 'non-configurational' language.)

(6) wror xo ba de ye lidəlay wi
brother indeed must your he seen is
'He must indeed have seen your brother'
(Tegey 1975:156)

5.3.2.1. Finnish *-kin*.

I chapter II, I described the syntactic placements for four sentential clitics. *-hAn, -kO, -pA* are second position clitics; *-kin/-kAAn* attaches to the finite verb. The sentential use of the clitic *-kin* is of interest here, because it constitutes a counterexample to Kaisse's claim that all sentential clitics must occur in second position.

Following the analysis given by Östman (1977), we can sort out three basic uses for *-kin*: (i) as a constituent modifier, it means roughly 'too, also', (ii) it can mark pure emphasis, and (iii) and it can be a sentential adverb -- i.e. it has a textual function. In this last function it always attaches to the finite

verb, and indicates either something unexpected, or if expected, then as something having been learned or something currently under discussion (Hakulinen and Karlsson 1979).

(7) Asetuksen antamista odotettiin, ja se annettiin-kin.
statute giving was awaited and it was given -also
'The giving of the statute was expected, and it was given, too.'

The local, constituent modifier use of *-kin* can be distinguished from its sentential use. In (8) it attaches to the main verb *ostanut* 'bought', and is a modifier of that word. In (9) it attaches to the finite verb, and is a sentential adverb.

(8) Kalle on ostanut-kin auton.
Karl has bought-also car
'Karl has also BOUGHT a car.'

(9) Kalle on-kin ostanut auton.
Karl has-also bought car
'Karl has also bought a car.'

In the case of simple verbs, there is a scope ambiguity. (10) can have either local or sentential scope.

(10) Kalle osti-kin auton.
Karl bought-also car
'Karl also bought a car.'

Finnish *-kin* is, then, a sentential clitic which does not occur in second position. Instead it attaches to the finite verb, which under standard GPSG treatments is the head of the sentence (i.e. S is a one or two bar-level expansion of V).

5.3.2.2. Veps *-(i)k*.

Somewhat similar facts obtain for Veps *-(i)k*, which can occur either in second position or after the finite verb (11-16). *-(i)k* is one of several interrogative markers in Veps, according to Kettunen (1943:532-34), on which I base this discussion. The others include *vn*, *jo*, *se*, and the Russian loanword *-l'i*. [I include Kettunen's Finnish translations in the following examples.]

(11) ortja jo joudab-ik magata?
Ortya EMP go -Q sleep
'Is Ortya going to sleep? Joutaakopa Ortja makaamaan?'

(12) ortja Ĭ-k vent pohjĭt' sina säigjärvhe?
Ortya not-Q taken bottoms there Säigjärvi
'Didn't Ortya take the (shoe) bottoms there to Säigjärvi?
Eikö Ortja vienyt kengänpohjia sinne Säigjärvelle?'

(13) defõnha mä voiń-ik mändä pimedas?
village I can -Q go dark(INES)
'Can I go to the village in the dark?
Voinko minä mennä kylään pimeässä?'

(14) mujõda mä sinum bokad roñt'iń-ik?
feel I your side dare -Q
'Do I dare feel your side?
Uskallanko minä tunnustella sinun kylkeäsi?'

(15) pahõń-ik mä söin?
much-Q I ate
'Did I eat much? Kovastiko minä söin?'

(16) küzü piga-ik hänen [saharan] tobad.
ask soon-Q it sugar bring
'Ask if they will bring it [the sugar] SOON.
Kysy, pianko he sitä [sokeria] tuovat.'

Examples (11-14) demonstrate the attachment to the verb, regardless of the positioning of that verb. The examples in (15) and (16) show *-(i)k* in second position. The other interrogative markers do not enjoy the same privilege of occurrence. *Jo* precedes the questioned constituent; it is not a clitic. Enclitic *se* occurs in a number of syntactic slots (as does its plural *ne*), predominantly second position or after the verb. *va* and the Russian borrowing *-l'i* are second position [semi-]clitics.

(17) keskĭń uńiš va voib noustta?
middle dream Q can wake up
'Can one wake up in the middle of a dream?
Kesken unta voiko kerättää?'

(18) jo se teile ĭl'end baifĭl' tütrĭd' ot'itä kazatšihan?
Q you not-was gentlemen girls took maid
'Really weren't there the gentlemen's girls for you, you took a servant girl? Tokkohan teille ei ollut (herrojen) herroilla tyttäriä, otitte piian?'

What is clear from Veps *-(i)k* (and the other interrogatives) is that it does not necessarily occur in clausal second position. It can instead attach to the finite verb of the clause. It is fairly clear that *-(i)k* is a sentential clitic, because it makes reference to the sentence as a whole for its positioning (and because it has the sentence as a whole in its scope). Finnish *-kO* and Veps *-(i)k* should be included among the large number of interrogative clitics and semi-clitics from languages all over the globe: Tagalog *ba*, Slavic *l'i*, Luiseño *-ɬu*, Dyirbal *-ma*, etc.

5.3.2.3. Permic-Ugric *-V*.

As a final counterexample to Kaisse's principle, I discuss the Permic-Ugric interrogative clitics. These have the same meaning as the Finnish and Veps interrogative clitics cited above. Hungarian *-e*, Ostyak *-ə*, Vogul *-ä*, Ziryene *-e*, and Votyak *-ə(ɯə)* are cognate morphemes. I assume as a starting point that these interrogative morphemes are functionally and semantically equivalent to the other sentential clitics of section 5.3.2.

These clitics do not occur in second position as Kaisse's principle predicts; rather they attach to the head of the predicate -- one cannot simply say verb here because in Hungarian, at least, present tense copula constructions are verbless, and the clitic attaches to the predicate adjective or predicate noun. Some examples of these clitics are listed below.

Hungarian (Koski and Mihafyly 1964:396):

(19) kívácsi volt, János-e as a gyerek.
curious was John-Q that the boy
'He was curious about whether that boy was John.'

(20) Vajon eljön-e Péter
whether comes-Q Peter
'(I wonder) whether Peter is coming.'

Ostyak (Fokos-Fuchs 1962:110):

(21) mant num-tûrum Xujan jina juXtupteta omat tâidan-a
me upper man water bring strength have
'Bist du imstande (hast du die Möglichkeit), mich zu den Männergewässern der oberen Welt zu bringen?'

Vogul (Fokos-Fuchs 1962:110):

(22) tit Xujew-ä?
here sleep
'Werden wir hier schlafen?'

Ziryene (Fokos-Fuchs 1962:111; Fuchs 1949:184):

(23) men poźe-e led't'äivni pristańe? -- poźe
me may-Q descend landingplace may
'Darf ich zum Hafen hinuntergehen -- Ja. Licet-ne mihi descendere? -- licet.'

Votyak (Fuchs 1962:184, 1937:319):

(24) kišno baśtom-a, um-a baśti
wife take -Q not-Q take
'Nehme ich (mir) wohl eine Frau, oder nehme ich (mir) wohl nicht?'

5.3.3. The Case Against Kaisse.

The above three examples (Finnish *-kin*, Veps *-(i)k*, and Permic-Ugric *-V*) demonstrate that sentential clitics are equally liable to attach to the head verb as to the first constituent of the sentence. And I have still more examples of this phenomenon, but three such examples are sufficient for our purposes. These are direct counterexamples to Kaisse's claim. Hence we are forced to reject the principle on empirical grounds.

Even on theoretical grounds her principle is suspect. It is not clear to me where in the grammar this statement should go. Kaisse apparently intends this to be a condition on or rule of cliticization; therefore it belongs in the clisis component. But I have argued in the preceding chapters that the gross syntactic positioning of clitics takes place in the syntax module, and not in the clisis component. Under my view, Kaisse's rule is a syntactic statement, of which I capture a part in my lexical feature [±B], ID feature [FIRST], and LP statement [FIRST] < [+B]. If it is to be considered a syntactic principle of ordering, then it constitutes a mixing of levels: this syntactic statement contains reference to clitic location. Following Zwicky and Pullum (forthcoming), we note that this addition of reference to non-syntactic items, i.e. clitics here, would be a weakening of the theory (cf. chapter VII).

In sum, we find Kaisse's statement concerning the location of sentential clitics of little use (except insofar as it characterizes a general tendency). It is not exceptionless, as she herself notes. Instead, we want to predict the locations of sentential clitics from the syntax of the language, especially the linear precedence statements. Syntactic positioning can be determined in this way for optional location, as was the case with Veps *-(i)k* above (5.3.2), where the clitic occurs either in second position or after the verb.

Hence I reject Kaisse's very strong claim, in favor of the heads/margins approach of Zwicky (1977) and Klavans (1982a), though I argue that these facts fall out from the syntactic description of a language (cf. chapter IV).

5.4. On the Phonology of Clitics.

Many analysts have noted that clitics always attach externally to hosts -- as enclitics or proclitics. Indeed these two types of clitics predominate in the literature. But there are a few examples of the phenomenon of **endoclisis**. Endoclisis is the phonological location of a clitic inside its host.

Most of the original examples of endoclisis cited by Zwicky (1977) have been reanalysed as something else. Thus although Zwicky cites Madurese, Estonian, Turkish, and Hua as languages exhibiting endoclitics, the data from these languages have been reanalysed as either as the effect of a morph metathesis rule (Zwicky and Pullum forthcoming) or as the result of the cliticization of an inflected word (Klavans 1980). In the case of Estonian, the 'endoclitic' *-gi* is revealed to be a lexicalized, derivational suffix, which is not connected synchronically to the [semi-]clitic *-gi* (Nevis 1984d, 1985a).

It follows from these results that surface endoclisis is always due to something else in the grammar (or lexicon), and should not be incorporated into the operations of a cliticization component. Liaison is the phonological attachment of one word to another, and not within another. Liaison does not include any operation akin to infixation.

5.5. Clitic Diachrony.

In this section I address two issues, the origin of second position enclisis and the Agglutination Hypothesis. Section 5.5.1 deals with how words and clitics come to be second position enclitics, section 5.5.2. with the Agglutination Hypothesis. My intent here is not to resolve any of the controversies, but to indicate how my bound word analysis of the Finnish particle clitics can bring evidence to bear on diachronic approaches to the phenomenon of Wackernagel's Law.

5.5.1. Second Position Enclisis.

Steele (1975, 1977), Hock (1982a,b), Comrie (1981), and Cowan (1984) are concerned with the tendency for auxiliaries and other sentential clitics to 'migrate' toward clausal second position, and the implications of this tendency as a mechanism of diachronic change.

Comrie (1981:86) says, "Second position is extremely rare as a required position for separate words", and therefore second position clitics should be excluded *ipso facto* from work on word order reconstruction, since they do not (necessarily) represent an earlier word order.

In this dissertation I have taken a different vantage point. Synchronic word order can indeed have recourse to such notions as '(clausal) second position'. In particular, the second position clitics of Finnish are very stable, dating back in some instances (*-pA, -kO*) to Proto-Balto-Finnic (circa 100 BC) or beyond. The 'category' of second position clitic has been enriched through the addition of *-hAn* (and in some dialects, also *-stA*). Similarly,

Proto-Uto-Aztecan can be reconstructed with second position elements (Steele 1981), which are continued in a majority of the descendant languages.

Following a suggestion by Steele (1977), both Hock (1982a,b) and Cowan (1984) use Second Position Enclisis as a motivation for constituent order change (SOV > SVO). Under their hypothesis, a (clitic) auxiliary verb shifts to second position for some unknown reason (other than Steele's 'universal tendency'), then other finite verbs follow suit, and finally, the pattern is generalized to include all verbs:

S O V > S=aux O V > S V O > S V O
[+fin] [-fin]

I am in sympathy with Steele's line of work. There is a tendency for members of certain semantic groups to occur at or near the beginning of a sentence. And since these semantic elements are frequently realized as verbal affixes and/or auxiliary verbs, this means that auxiliary verbs will enjoy an initial positioning (or close by).

If clitics are linguistic primes (as suggested by Steele 1976), there is no logical connection between the notion 'clitic' and other word classes, such as auxiliary verb. Furthermore, there is no motivation for any generalization from second position clitic to finite verb. But, if second position clitics are members of word classes, such as auxiliary verb, pronoun, adverb, and the like, then we can easily account for the extension of a pattern from (clitic) auxiliary verb to finite verb and ultimately to all verbs.

And since second position is available to full as well as bound words, Hock's (1982a) approach need not rely on the clitic status of auxiliary verbs (though in fact he takes special pains to point out the clitic status of Second position auxiliary verbs in Germanic, Romance, and Kashmiri).

5.5.2. On the Agglutination Hypothesis.

My bound word/phrasal affix analysis has implications for the agglutination cycle in language change. Here I dwell on one small aspect of the theory, as the topic goes well beyond the bounds of this dissertation. The point to be made here is that second position (sentential) bound words never complete the agglutination cycle by becoming second position sentential affixes.

This observation will have implications for the reconstruction of word order on the basis of morphology (as in Givón's (1970) slogan, "Today's morphology is yesterday's syntax"). What I claim here is that the Finnish second position clitics are bound words. An extension of this analysis will be the much stronger claim that all sentential second position clitics are bound words. In fact, I know of no true second position affixes. In the more carefully

studied languages having second position clitics, i.e. Finnish, Luiseño (Steele, 1976, 1977,; Steele et al. 1981), Lummi (Steele et al. 1981), Latin (Janson 1976), and Warlpiri (Hale 1973), the second position clitics are all bound words. Such facts suggest that second position bound words are unlikely to complete the agglutination cycle. Thus bound words are not likely to become sentential affixes in sentence-second position. These clitics are not really part of word morphology, and do not contribute to the methodology of predicting a former word order within a sentence from a morpheme order in a word.

In particular I have have in mind some research on the development of the Balto-Finnic second position clitics *#-pa*, *#-s*, and *#-ko* in Old Estonian. In Finnish two of these clitics, namely interrogative *-kO* and emphatic *-pA*, are clearly bound words -- as demonstrated in chapters II and IV. Former clitic *-s* has become lexicalized, cf. 2.1.2.1. In Estonian the morphemes cognate to Finnish *-s* and *-pA* are no longer bound words, but are full words, and in one instance have achieved some extra syntactic freedom in sentential positioning. Here we are dealing with the phenomenon of decliticization -- a former clitic takes on phonological independence to become a full word (Jeffers and Zwicky 1980:57-58). In the case of *#-ko*, the morpheme has not become a sentential affix in Estonian, but has disappeared entirely, leaving only a few relics in combination with other morphemes.

The data have been described in some detail by Ariste (1973) and Alvre (1976, 1981) -- all three works in Estonian (but for summary and analysis, see Nevis 1984a, to appear). I summarize their results here. Estonian underwent apocope, as well as numerous other sound changes in its development from Proto-Balto-Finnic. In sentence-initial words to which either *#-ko* or *#-pa* was attached, the final vowel of the bound word (i.e. *-o* and *-a*) was apocopated. In other words, where there was no bound word, some other vowel was dropped -- in many instances this vowel was *a* or *ä* (depending on vowel harmony), because those two vowels occur most frequently in case endings. The vowels of the case endings were not apocopated before the former *#-pa* and *#-ko* enclitic words (which thus acted prophylactically). As a result, the vowels occurred only before *-p* and *-s*, and were then reanalysed as part of the *-p* and *-s* morphemes. Furthermore, in certain non-initial syllables, *ä* became *e*, so that *ep/äp* and *es/äs* were the outcome.

These two words underwent leveling in favor of the *e*-variants, yet retained their syntactic positioning in the second slot of the clause:

(25) Nüüd es tee uskute (from 1686 New Testament, cf. Ojansuu 1922)
now Q you believe
'Now do you believe?'

(26) selle kivi peal ep kolgitigi neid riideid
this stone on EMP one-pounded these clothes
'On this stone one pounded these clothes.'
(from Saareste 1958, see also Ariste 1973:33)

This Old Estonian instance is really more complicated than mere decliticization because there was at the same time morphological split. Some *ep* and *es* morphemes became separate words; others became affix-like endings to existing words.

(27) eks 'isn't that so' < *ei-kö-s (ei = negative verb, -kö = Q)
ons 'is it?' < (on = 'is')

(28) see'p se on 'that's that' < see-pä see on

Estonian *es* and *ep* are full full words having clitics as their etymological sources. Under my bound word analysis the only major change to take place (other than the reanalysis of the morpheme boundary before the stem vowel) was the loss of the liaison rule from the grammar. Estonian seems to be typical with regard to the diachronic behavior of second position bound words: the morphemes in question do not complete the agglutination cycle in becoming affixes. My discussion of dialectal *ko* in the Finnish dialects (in chapter II) demonstrates that the second position clitics can swing back and forth between bound word and semi-clitic (and perhaps even full word).

By the same token, phrasal affixes are not likely to be reinterpreted as words. That is, they will not decliticize so readily as bound words.

5.6. Summary.

Many of the ideas presented in previous chapters (II-IV) are refinements of analyses already existing in the literature. The idea that words which undergo liaison to become bound words are marked in the lexicon with a feature [±liaison] is due to Klavans' (1982a) feature [±clitic]. My binary division of clitics into two types is somewhat similar to that of Zwicky and Pullum (1983), though the border is drawn differently. The independence of the phonological and syntactic parameters is most strongly defended by Klavans (1985), and in applying this distinction I have also drawn on Kaisse (1985).

In chapter VII, the conclusion, I examine a modular scheme for grammar proposed by Zwicky and Pullum (forthcoming), Zwicky (1982a), and Kaisse (1985). I incorporate my bound word analysis into their interface program, and on the basis of that, revise their componential divisions. The result is an autonomous readjustment component, which is highly restricted in power.

Endnotes.

1. Notably missing from this list of prominent Finnish linguists is the name of Paavo Siro. Siro, in his (1977) book Sijakielioppi (= Case Grammar), does not mention the particle clitics. Case grammar focuses on accounting for 'deeper' semanto-syntactic grammatical relations, and thus ignores the 'surface' syntactic positioning of sentential adverbs, conjunctions, and sentential clitics necessary for the description of sentential clitics.

2. As I noted in my LSA article (Nevis 1984), there are some factual problems with this account. First, *hindi* is not a true clitic, instead it acts as host for the second position "clitics" of Tagalog. Second, the second position morphemes of Tagalog appear to be semi-clitics, not true clitics. They are strictly positioned in second position, and are unstressed, but these two facts are insufficient to establish clitichood. Even Wackernagel (1892) noted this in his use of the term 'Quasi-Enklitikon' -- an unstressed word in sentential second position.

3. Zwicky (1984a) argues that the German, Chrau, and Hidatsa clitics are actually not clitics at all. They are all shown to be something else -- full words or full affixes, depending on the language. So Kaisse's data are in fact not counterexamples to her own claim. However, the Finno-Ugric data are an embarrassment to Kaisse's principle.

4. This corpus, collected for a 1984 seminar paper, includes the 48 languages below, arranged by family (rather than phylum) -- according to the classification of Voegelin and Voegelin (1977). The languages on this list exhibit at least one Wackernagel-type clitic, though I have not distinguished clitics from semi-clitics here.

Algonkian: Cree, Fox Abnaki (= Penobscot), Ojibwe
Arawakan: Parecis
Athabascan: Navaho
Balto-Finnic: Finnish, Veps, Votic, Karelian
Baltic: Latvian, Lithuanian
Costanoan: Mutsun
Hellenic: Ancient Greek
Italic: Latin
Kwakiutlan: Kwakwala
Lappic: Ruija Lapp, Northern Lapp
Mayan: Tojolabal Maya
Miwok: Northern Sierra Miwok
Ngarga: Warlpiri
Nootkan: Nitinaht
Numic (= Plateau Shoshonean): Southern Paiute, Mono, Northern Paiute, Shoshoni

Paaa-Nyunga: Warumungu, Ngiyambaa, Dyirbal
Pamir: Pashto
Pomo: Eastern Pomo
Sahaptin-Nez Perce: Sahaptin
(Coast) Salish: Lummi, Halkomelem, Bella Coola
Siouan: Crow
Slavic: Russian, Bulgarian, Serbo-Croatian, Czech
Sonoran: Pima
Tagalic: Tagalog
Takic (= Southern California Shoshonean): Cahuilla, Cupeño, Luiseño, Serrano
Tübatulabal: Tübatulabal
Wiyot: Wiyot

All of the languages on this list display a large degree of free constituent order, and some have free word order as well. Notably absent from this list are languages, like English, having strict ordering of elements. Deviations from totally free constituent order are verb-initial languages like Tagalog, Lummi, Nitinaht, and a few others. Note that the verb-initial languages also have fairly free constituent order other than the initially positioned verb; complements of the verb in these languages are not so strictly ordered. Even so, a certain number of elements can precede the verb (mostly words meaning 'not', or acting as adverbs, conjunctions, and topicalized nominals), in which case the Wackernagel-type clitics precede the verb and attach to whatever occurs in initial position.

CHAPTER VI

LANGUAGE EXTERNAL EVIDENCE FOR CLITICS AS WORDS:
LAPPISH PARTICLE CLITICS

6.0. Introduction.

In the preceding chapters I have argued for the position that the Finnish particle clitics, and indeed all second position clitics, are **bound words** rather than phrasal affixes or linguistic primes. I motivated this position with synchronic language internal facts. In arguing a particular stance on the word/affix status of clitics, linguists have, in general, focussed on language internal evidence -- formal aspects of these clitics' synchronic syntax/morphonology -- to the exclusion of language external evidence (other than the diachronic developments of clitics). In this chapter, then, I examine some external evidence, showing that that the evidence supports a bound word analysis over a phrasal affix.

Let me review some assumptions and arguments from chapter II. I assume that syntax refers only to the binary distinction between **words** and **affixes** and that the latter are represented as feature complexes (cf. chapters two and four). In addition, I argue that clitics do not constitute a **word class**, but encompass other word classes, such as adverb, pronoun, and "any word class ... save (lexical) verb and (lexical) noun" (Klavans 1979:36). In addition, I take the position, argued by Zwicky (1984a), that all words are to be assigned a word class, and that there are no acategorematic words. It is true, however, that it is sometimes difficult for an analyst to assign word classes, especially in the absence of inflectional morphology on words (as is often the case for clitics).

6.1. Evidence.

I turn now to types of evidence used for argumentation in the subfield of cliticization. Most of those works cited above rely on (language-internal) synchronic evidence, as do many other studies: e.g. Hale 1973; Steele 1977; Steele et al. 1981; Kaisse 1981,1982; Klavans 1980,1982a,1983,1985; Smith and Johnson 1984; etc. Often in

conjunction with the synchronic accounts are analyses of the diachronic developments of clitics. These tend to focus on how or why sentential clitics migrate to sentence-second position (see, for example, Comrie 1980; Steele 1976; Hale 1973) or on the interactions between Wackernagel's Law and basic word order (Hock 1982; Comrie 1981; Cowan 1984; etc.), cf. chapter 5.5. That is to say, examinations of clitic phenomena have been restricted to synchronic and diachronic evidence. Other kinds of language-external evidence, however, should be as revealing in the domain of cliticization as it has been in phonology (proper).

6.2. External Evidence.

Ideally we want to reinforce argumentation with evidence from several source-types. For this I turn to external evidence; some of the external evidence used in the subfield of phonology may be applicable here: child language acquisition, aphasia, orthography, borrowing, language play, speech errors, and the like.

Orthography will be extremely difficult to evaluate. Clitics are sometimes written together with the host, sometimes not. Sometimes they are written together with each other (without being attached orthographically to the host), sometimes not. In many instances the written representation of clitics continues a stage of the language at which the clitic was a free form.

In Lappish, for instance, various orthographies (from Korhonen 1981: Nielsen, Friis-Bergland-Ruong, Ravila-Itkonen, Arjeplog Language Committee) represent the clitics in one of two ways -- as free words or as morphemes bound to the host. To a certain extent orthographic differences coincide with the Finnish-Norwegian national boundary. On the Finnish side, in the Ravila-Itkonen system the clitics are written attached to the host as in Standard Finnish; on the Norwegian side, the Nielsen system captures them as independent words.

Child Language Acquisition. Children acquire clitics according to the function and meaning of the morpheme rather than word or affix status. In the acquisition of clitic morphemes, function apparently overrides form. This is because the content that clitics carry are equally likely to occur in full words or in inflectional affixes. Thus, Toivainen (1984) observes that of the Finnish second position clitics interrogative *-ko/kö* is acquired rather early, along with certain interrogative pronouns. (Other clitics, such as Finnish *-han/hän*, the marker of shared information between the speaker and hearer, make a rather late appearance, due to their vague meaning and low functional load.) Similarly T. Itkonen (1981) claims that the emergence of inflectional affixes in children's speech takes place at the same time as the emergence of grammatical ("function") words. In light

of this, I see no means to derive from the realm of child language acquisition any evidence that would bear on the word/affix status of clitics.

Studies of the behaviors of clitics in language play and aphasia would be most interesting and most welcome. Unfortunately I am not aware of any relevant studies. We would want, of course, to determine whether clitics pattern with full words or with inflectional affixes.

6.3. Evidence From Borrowing.

I focus now on the evidence from observations on the borrowing of second position clitics from one language to another. So far as I know, no such evidence has been presented in the literature. Therefore I present some data from Finnish borrowings in Northern Lappish to demonstrate how this type of external evidence can be of use.

First, some general remarks on borrowing: We know from numerous sources (see Thomason 1978 for references) that words are most frequently borrowed, inflectional affixes by comparison very infrequently. As an implicational statement, we can say that if there are inflectional affixes borrowed from one language to another, there will also be loanwords. The reverse, however, does not hold true: if there are loanwords in a contact situation, inflectional morphology need not also be borrowed.

Following this line of reasoning, then, if we find loanwords, we do not necessarily expect to uncover loaned morphology. It is this situation I argue to exist in northern Scandinavia. The Lappish languages exhibit large amounts of borrowed vocabulary from Finnish, but little or no morphological interference. Alongside the loanwords are several second position clitics.

Before I present a detailed description of the Lappish situation, I offer two other parallel instances of the borrowing of second position clitics in northern Europe:

a. The Balto-Finnic clitic *-pa/-pä* (which marks emphasis) has counterparts in the Indo-European Baltic languages -- Latvian *ba* and Lithuanian *bə̀, bə́*, both apparently with the same meaning. L. Hakulinen (1979) suggests that the direction of borrowing has proceeded from Proto-Baltic to Proto-Balto-Finnic along with numerous other loanwords. Ariste (1973), however, suggests the opposite direction of borrowing (PBF > PB) since he connects Balto-Finnic *-pa* with Permic (Komi *pö*, Udmurt *pe*). This situation resembles the Finnish-Lappish one to be described below.

b. The language contact area in northeastern Russia exhibits several second position clitics. Lydic and Veps possess, besides the native inventory of particle clitics (i.e. *#ko, #pa, #s*, as in Finnish), also several Russian loans. Larjavaara (1979) mentions, for example, *ved', -že (-žo)*, and, as a loan-translation, Russian *to* > Lydic, Veps *se*. (The use of these particle clitics in Russian is not known in the western and southern dialects, but is wide-spread in the eastern and northern dialects; Larjavaara 1979:125, Kuznetsov 1960:125-126). One finds in Veps also the Russian loan *-l'i*. Again, this situation is like that of the Finnish-Lappish situation discussed in more detail below.

Due to space considerations I cannot evaluate these two contact situations here, but I cite them as suggestive evidence that second position (and, generally, sentential) clitics are easily borrowed, and are therefore to be included with the class of loanwords, rather than (the much more rarely borrowed) inflectional morphology. I will, however, present in the following a description of the Finnish-Lappish contact situation that supports my claim.

6.3.1. The Finnish-Lappish Contact Situation.

The origins of the Lappish language group have been open to much speculation. It is on the surface the most closely related (Finno-Ugric) branch to Finnish (that is, outside the Balto-Finnic sister languages). The Lappish proto-language, "Common Lappish" (circa 1000-700 B.C.), can be reconstructed and appears to be very similar to "Late Proto-Finnic" (i.e. Balto-Finnic, circa 1000-0 BC), both of which derive from "Early Proto-Finnic" (1500-1000 BC) (Korhonen 1981:27). The binary branching of Early PF has been attributed by some to language contact -- language shift from some hypothesized "Proto-Lappish" to Proto-Finnic (see Korhonen 1981 for a discussion), but this hypothesis remains controversial.

The time depth that I address here is nonetheless more recent than the Balto-Finnic/Lappish split (3000-3500 years ago). For example, the sound change *#š>s* in Lappish yields the third person singular pronoun *son*, but *#š>h* in Finnish yields the same pronoun as *hän*. The Finnish pronoun also cliticizes: *-han/hän* (distinct in meaning and syntax from its pronominal source). Lappish has two corresponding clitics -- *-son* and *-han* (both with the same meaning as Finnish *-hAn*). The latter is clearly a loan from Finnish due to the presence of the *h* instead of native *s*; and the former is suggested to be a loan translation.

Many of these loans are quite old, though, as they appear in all the Lappish languages, not just Northern Lappish. Thus, we note, the adstratum situation has continued for numerous centuries. This is not, by the way, an intensive contact nor one of language

shift -- one can cite in this regard the existence of Lappish-to-Finnish (Finnish dialects and Karelian) loanwords (Korhonen 1981:39-40). This Lappish-Finnish situation involves bidirectional cultural influence, rather than unidirectional pressure on the part of the Finnish speakers.

In a typical borrowing situation, the first foreign elements to be incorporated are words (Thomason 1978, 1980). If "there is strong cultural pressure from the source-language speakers on the borrowing-language speaker group then structural features may be borrowed as well -- phonological, phonetic, syntactic, and even (though much more rarely) features of the inflectional morphology" (Thomason 1978:8). But the situation in northern Scandinavia apparently lacks this strong cultural pressure on the part of the Finnish speakers. The situation exhibits by far mostly loanwords, little or no phonological-phonetic borrowing[1], and perhaps some borrowed syntactic structures (e.g. the predicative copula *lse-*, several coordination and subordination constructions, and possibly also SVO constituent order; see Korhonen 1981:342-6). Clearly the situation has not been so intense as to permit the (typically rare) borrowing of inflectional morphology.

A word on this last point: Morphologically (and typologically) Lappish has wandered far from the agglutinative parent language -- of all the Finno-Ugric languages modern Lappish tends most toward the flectional, or fusional- symbolic (Korhonen 1974, 1981:203). Balto-Finnic, by comparison, remains fairly agglutinative in nature.

To support my claim that the inflectional morphology of Finnish has not been incorporated into that of Lappish, I present in the appendix an overview of the two systems. It becomes clear from the presentation in the appendix that the Lappish inflectional categories and morphemes have arisen through genetic inheritance rather than interference from Finnish. (Since we are dealing with syntactic cliticization, I treat here only inflectional morphology and ignore derivational morphology.)

It is now established that Lappish evidences numerous loanwords of Finnish origin, but no interference from Finnish inflectional morphology. We arrive now at the main topic of this chapter: the borrowed particle clitics.

6.3.2. Finnish and Lappish Particle Clitics.

The meanings of the Finnish and Lappish particle clitics are strikingly similar:

Finnish	Lappish[2]	Meanings
-han/hän	hån, son	shared information between speaker and hearer
-pa/pä	bå (be)	emphasis
-kä	ge	'and' (in conjunction with the negative verb)
-ko/kö	go	marks a yes/no question, and indirect questions (the 'if, whether' subordinated questions)
-kin/ -kaan/kään	ge	'also, even'

The clitic status of the Finnish morphemes *-han/hän, -pa/pä, -kä, -ko/kö,* and *-kin/kaan/kään* has already been motivated in chapter 2. As a review I summarize the main arguments for considering these to be clitics. These morphemes behave like words in three ways -- they condition external sandhi in three rules: (1) stem allomorphy, (2) word-final *t*-assimilation and (3) word-initial gemination.

Word-final t-assimilation applies only across words as in (2a), never within words, as in (2b). The particle clitics permit the application of this *t*-assimilation rule as in (2c), thus behaving like full words. **Word-initial gemination** takes place between adjacent words, but never between a stem and its affix. It is morphophonemic (or 'morphophonetic' since it affects aphonematic ['] and [?] as well) and will geminate the initial consonant of a word following certain morphemes (often indicated diacritically by a raised /^{X}/ as in vene^{X} 'boat', mennä^{X} 'to go', tule^{X} 'come!', tietysti^{X} 'of course', talonsa^{X} 'his/her/their house'). For example, in (3a) the final /^{X}/ of vene causes gemination of the /t/ in tulee. (On the subject of word-initial gemination, see Karlsson (1983:348ff) and references therein.) In this respect, the particle clitics behave like independent words. We find [venek:in, men:äk:ö, tulep:a, tietüstik:in, talonsah:an] rather than *[venekin, men:äkö, tulepa, tietüstikin, talonsahan].

(1) **stem allomorphy**
nominative singular: hammas 'tooth'
stem: hampaa- (e.g. genitive hampaa-n, plural hampaa-t)
with particle clitic: hammas-han -- *hampaa-han

(2) **word-final *t*-assimilation** (optional)
a. miehet kuolivat [miehek:uolivat] 'the men died'
b. pitkä [pitkä -- *pik:ä] 'long'
c. en tiennyt-kään [en tien:ük:ä:n] 'I didn't even know'

(3) **word-initial gemination**
a. vene tulee [venet:ule: -- *venetule:] 'the boat comes'
b. vene-kin [venek:in -- *venekin] 'also the boat'

On the other hand, these morphemes are in some ways affixal: they are bound morphemes, which interact with their host prosodically as in (4) and, in a limited way, also phonologically (i.e. vowel harmony yields *perunahan*, not **perunahän*, and *minähän*, not **minahan*). The prosody rule in (4) operates at the word level, and is not to be confused with prosodic operations at the phrase level.

(4) a. péruna (*pérunà) kásvaa 'the potato grows'
b. pérunàhan (*pérunahan) 'the potato, you know'

This mixed status (word-like in syntactic and morphological properties, loosely affixal in phonological properties) is the profile of the bound word type of clitic, cf. chapter 2.

As for the Lappish particle clitics, one can say that there is less phonological and prosodic interaction with the host than exists in Finnish. So far as I know, no word-internal phonological rule (such as vowel harmony in Finnish) applies between the host word and the clitic nor does the morpheme interact prosodically with its host (5), as it does in Finnish (4a,b). The prosody facts for Lappish words are similar to those for Finnish: primary stress on the initial syllable, secondary stress on subsequent odd syllables, with the exception of the word-final syllable. Sammallahti (1977:94ff.) makes clear the similarities between the prosodic behavior of the particle clitics and the stress pattern of compounds; thus the facts indicate a word-like status for the morphemes.

(5) a. [táhka:pa -- *táhka:pà] 'does, makes (3DU)'
b. [táhka:lèa'pa] 'does, makes (2DU)'
c. [táhka:pako -- *táhka:pàko] 'does, makes?' (with question particle *go*)
(Sammallahti 1977:92-93)

The Lappish particle clitic, then, is not a true bound word, but illustrates the related phenomenon of semi-clisis (cf. chapter 4.1). That is, the Lappish "clitics" are prosodically, but not phonologically, subordinated to a neighboring word, and cannot stand in isolation or take phrasal accent. The semi-clitic class consists merely of stressless words in the same sentential positions as sentential clitics. Zwicky (1982a) refers to this class of prosodically subordinated words as leaners.

This difference between the Finnish bound word and the Lappish semi-clitic will not affect my argument below -- the morphemes are positioned in the same sentential slots (either in sentence second-position, as in (6), or after the finite verb, as in (7), depending on the morpheme); they have the same morpheme order when more than one occurs in the same slot (for example, Finnish allows *-pa-han* as does Lappish *-bâ-hân*); they are subject to the same cooccurrence restrictions; and they have the same meanings.

(6) Finnish: Hyvää-pä-hän siitä tuli.
good-EMP-HAN from-it came
'Some good came from it, you know.'
Lappish: buorrĕ-bâ-hân dâst bodii.
good-EMP-HAN from-it came
'det blev da noget godt av det.'
(Nielsen 1979:104)

(7) Finnish: Minä en tarkoita-kaan tätä perintöä!
I not mean-EVEN this inheritance
'I don't even mean this inheritance!'
Lappish: mon im oaiveld-ge dâm arbe.
I not mean-EVEN this inheritance
'Ich meine auch nicht diese Erbschaft.'
(E. Itkonen 1960:110)

The Finnish clitics *-han/hän, -pa/pä, -kä, -ko/kö,* and *-kin/kaan/kään* have been borrowed by Lappish speakers as the semi-clitics *han/son, be/ba, ge, go, and ge.* Many of these morphemes have grammatical function or subtle meaning, perhaps indicative of inflectional affixal status. However, they do fit among the numerous conjunctions and subordinators borrowed from Finnish into Lappish (e.g. *jâ < ja 'and', dâhjĕ < tahi 'or', vai < vai 'or', muttĕ < mutta 'but', âttĕ < että 'that', mâdĕ ... dâdĕ < mitä ... sitä 'the more .. the more', jos < jos, vaikĕ, vaikĕ < vaikka 'although'* -- Korhonen 1981:346).

There are two potential counterarguments to my analysis. First, there might have been reanalysis of the morphemes during the borrowing from Finnish to Lappish; and second, the borrowed particle clitics might indeed be inflectional morphology, only no other part of inflectional morphology has as yet been borrowed.

It is clear that the Lappish speakers have not borrowed the Finnish particle clitics as clitics in the strict sense of the term. So, one could argue that there has been a reinterpretation of original (putative) phrasal affixes as prosodic leaners. (Phrasal affixes are, after all, the loosest of inflectional morphology.) But this would not enable us to capture the fact that sentential positioning and morpheme order are the same in the two languages, as are their cooccurrence restrictions. If the Finnish "phrasal affix" had been reinterpreted as some sort of prosodically weak word, why would the new word have the same distribution as a sentential clitic? It is at least clear that the prosodic subordination of the morphemes has been borrowed along with the morphemes themselves.

Another potential problem crops up for my claim that the particle clitics pattern in this borrowing situation with the loanwords rather than with the (non-existent) loaned morphology: even though no other inflectional affixes have been borrowed between the two languages, it could be that the particle

clitics are the leading edge of borrowed inflectional morphology. Again, this is plausible because of the looseness of the attachment of this class of morphemes. However, given the time depth that we are dealing with, it is rather peculiar that no other inflectional morphology has been borrowed in the centuries that have passed. Furthermore, as noted above, the contact situation is not intensive enough for the borrowing of inflectional morphology.

6.4. Summary.

I have argued that the phonologically bound word in Finnish corresponds in nearly every parameter to the borrowed Northern Lappish prosodic leaner -- the only difference is phonological interaction with the host word. The fact that the Finnish clitics are so easily borrowed in such an adstratum situation as this constitutes evidence that the clitics in question are bound words, because the loaned clitics pattern with the loaned words rather than with non-existent loaned (inflectional) morphology. Bound words are more easily borrowed under casual contact conditions than are phrasal affixes because words are more readily borrowed than inflectional affixes.

Endnotes

1. There are, in fact, striking phonetic-phonological differences between the two languages beyond the divergent phonemic inventories. Northern Lappish has three degrees of length, Finnish has only two. Lappish has many more consonants but fewer vowels than Finnish. Finnish has vowel harmony -- lacking in Lappish. Lappish has preaspiration, which is missing in Finnish . Although both languages have Consonant Gradation, CG in Finnish is primarily phonological; in Lappish it is entirely morphological.

2. The list here is fairly complete for the Finnish particle clitics -- only *-s* is missing. I argued in chapter 2 that *-s* is not a true clitic, but more affixal in nature. It is of interest to note that this more affixal morpheme has not made its way from Finnish to Lappish. I take this as a justification of my analysis in chapter 2, in which I distinguished between the word-like particle clitics *-hAn*, *-kO*, *-pA*, and *-kin/-kAAn* on the one hand, and affixal *-s* in the other.

The list for the Lappish particle clitics is not complete. Here are some other second position clitics in Northern Lappish: (from E. Itkonen 1960)

-nâi 'auch, sogar' (cf. *-ge*, both occur in the same slot in the sentence and word)

-râk, -râkkâ(i)n an emphatic particle clitic

Collinder (1949:228) mentions also a *-ki,-k* variant in place of *-ge*, as well as an interrogative *-pai* (in questions expecting the answer 'no') E. Itkonen (1969:110) lists *-bai* as a variant of *-ba*, noting that it can be used in interrogative function rather than emphatic function.

The term 'Lappish', by the way, seems to be falling into disuse these days (since apparently it can have offensive connotation in the Scandinavian languages). The native term 'Saame' is preferred now in Lappology. I retain the appellation 'Lappish' as a correspondent of the Finnish word 'lappi' because it will be much more familiar to the majority of English speakers than the comparatively unfamiliar 'Saame'.

CHAPTER VII

CONCLUSION

7.0. Introduction.

In the preceding chapters I have detailed a proposal to treat the Finnish particle clitics as bound words. I have shown that the particle clitics pattern syntactically and prosodically with class [41] sentential adverbs. The clitics *-hAn*, *-pA*, and *-kO* have the same distribution as adverb subclass [41.6]; clitic *-kin/-kAAn* has the same distribution as subclass [41.8]. Furthermore, members of this adverb class typically do not bear any sentential stress; and in particular, at least one member, *sitä*, cannot bear phrasal stress. Although inherent stresslessness is not equivalent to (not sufficient for) clitichood, the two notions are quite similar. That is, both clitic *-hAn* and semi-clitic *sitä* are unstressed (syntactic) words; the difference between them is the application of a rule of liaison to *-hAn*.

My goal has been to demonstrate that sentential and second position clitics such as those found in Finnish are not to be considered representative of a linguistic prime, clitic; they are instead syntactic words that have lost their status as phonological words. Nor are they to be considered phrasal affixes, as demonstrated in chapters two and six. In chapter two I showed on language-internal, synchronic grounds, that the Finnish particle clitics have very few characteristics of inflectional affixes, but instead have the syntactic and even morphological qualities of independent words. And in chapter six I presented some language-external evidence from the borrowings of particle clitics (from Finnish into Lappish) that indicates a word analysis of the particle clitics is to be preferred over an affixal analysis. Section 5.5.2. also mentions some diachronic evidence.

In this final chapter, I offer a model of module interaction in the grammar; a model fully consistent with Zwicky and Pullum's Interface program, which entails Zwicky's (1970) Principle of Phonology-Free Syntax (henceforth, PPFS) and Zwicky's (1969) Principle of Superficial Constraints in Phonology (PSCP, as

instantiated by Kaisse 1985). My approach is also consistent with Klavans' (1985) arguments to the effect that the syntactic and phonological parameters of clitics work independently of one another. But, contra Kaisse 1985, I do not separate special cliticization and simple cliticization into distinct components. Instead I offer a highly restrictive, unified theory of liaison (which is to be kept separate from phrasal affixation).

Having purged the cliticization component of its syntactic powers, I can now reduce the notion 'cliticization' to mere liaison -- phonological subordination (or readjustment) of a word to some host. Syntactic conditions, such as those referring to edge of constituent, head of constituent, and c-command (pointed out by Kaisse 1985), are permitted (cf. section 5.3), but syntactic operations, such as immediate dominance and linear precedence, are not permitted in this component. This brings cliticization more in line with other operations in this readjustment component -- external sandhi, flattening, and the like (see below). Liaison and external sandhi are phonological operations with strikingly similar syntactic conditions on their applications.

7.1. Autonomous Components.

Zwicky and Pullum (to appear) offer a theory of grammar with limited expressive power and high modularity. In their framework the notion grammar is a composite of autonomous modules, which interface linearly. The 'Strict Autonomy Hypothesis' requires that grammars have, at least, four properties discussed by Sadock (1983b) and Zwicky (1984c): nonuniformity, limited interaction, sharpness of boundaries, and nonredundancy of function.

First, separate components typically exhibit nonuniformity. That is, they are formally distinct pieces of the grammar. The operations in one component should not have properties in common with the operations in another component.

Now, Zwicky (1984c:367) argues that "Nonuniformity alone is scarcely enough to warrant a division between components." Instead:

> True uniformity would make a strong case for a single component embracing both the principles in [component] 1 and [component] 2. (Zwicky 1984c:367)

Second, separate components interact only in limited ways; the rules and conditions governing one component have limited or even no access to those rules and conditions governing another. A metalinguistic principle restricting potential interactions here is the PPFS, whereby the principles regulating the units and structures of the syntactic module have no access to phonological information, cf. section 7.1.1. This generalization follows from a strict separation of the syntactic and phonological modules.

However, in the linearly organized model that I assume here, components that are fed by preceding components do not interact with preceding components. This is what Zwicky calls 'backward interaction'. With truly limited interfacing, even such backward interaction is rejected in favor of that which is restricted to output information of a preceding component. Thus, the interface program does not permit information from the input or interior of another component, and global rules are disallowed.

One limitation on interaction is enforced by the PSCP, according to which the only syntactic information available to phonological rules is that which is represented in surface structure, cf. section 7.1.2.

Third, if modular boundaries are sharp, there should be no rules having a status intermediate between two autonomous components, i.e. rules having some properties of one module, some of the other. Dressler (1985) argues for such fuzzy boundaries, though with the stipulation that transitions between modules are often steep, rather than gradual. Highly relevant to our purposes here is Sadock's (1983a) clitic cline, which supposedly constitutes a fuzzy transition from syntax to clisis. I address this issue below in section 7.1.3.

The transition between cliticization and syntax is not fuzzy. I argue below that cliticization is autonomous with respect to the syntactic module. For now, let it be noted that Zwicky (1982b, 1984c) offers the observation that cliticization operations, which have sometimes been claimed to constitute rules intermediate between syntax and phonology, are formally distinct from both syntactic operations and other phonological operations. In particular, no cliticization rule ever feeds or bleeds a syntactic rule. Yet syntactic rules do feed and bleed cliticization rules.

Fourth, autonomous modules are nonredundant; they do not duplicate one another's functions or principles from other modules. This principle of autonomy is subject to some weakening, as suggested by Sadock (1983a)[1] and Zwicky (1984c).

In the instance of the cliticization/syntax interface, I believe that the two modules in question are nearly completely autonomous -- there is little uniformity, no interaction, sharp boundaries, and no redundancy of function.

One challenge to my claim of nonuniformity is Kaisse's (1985) observation that rules of cliticization and rules of local syntax (à la Emonds 1976) share reference to notions like 'c-command' and 'margin of constituent'. Of course, GPSG makes no corresponding distinction between local and nonlocal rules, so the extension of her analogy to GPSG syntax will not work here.

7.1.1. The Principle of Phonology-Free Syntax.

One restriction on the operation of syntactic rules is Zwicky's (1969) Principle of Phonology-Free Syntax. According to Zwicky (1969, 1982, 1984c) and Zwicky and Pullum (1984, to appear), the syntax module does not have recourse to phonology. This means that rules of syntax will not interact with rules of phonology. In other words, phonological rules will never feed or bleed (or counterfeed or counterbleed) syntactic rules. In addition, phonological information is excluded from playing any role in the operation of syntactic rules.

One set of putative counterexamples to this principle is the class of cliticization rules. An early view was that clisis rules are a mixture of syntactic and phonological operations. Generally these rules apply to words or morphemes, yet have phonological consequences. In the case of second position enclisis, it seems that a cliticization rule locates material after the first constituent or word of a clause -- a syntactic operation -- and at the same time attaches that material to some neighboring word -- a phonological operation.

In order to maintain the PPFS, Zwicky (1982b, 1984c) separates rules of syntax from rules of clisis, with the former preceding the latter. Zwicky argues that there are no clear instances in which a cliticization rule ever feeds or bleeds a syntactic rule, and that there are many instances in which a syntactic rule feeds or bleeds a cliticization rule. Therefore, all of syntax precedes all of clisis.

Bissantz (1985) claims that the monostratal GPSG framework actually requires a separate treatment for cliticization. She demonstrates that the syntactic facts for English auxiliary reduction and complementizer contraction are predicted by GPSG syntax, and that the framework forces liaison to constitute a distinct component.

A similar separation of syntax and clisis is offered by Smith and Johnson (1984).

7.1.2. The Principle of Superficial Constraints in Phonology.

As a component separate from syntax, cliticization now appears more phonological in nature. I argued in chapter 4 that the only operation a cliticization component can have is the readjustment operation of liaison (see also 7.1.3 below). Should this approach prove too strong, one must still determine how much syntax is available to the cliticization component.

I turn now to the PSCP (which is the same as Kaisse's 1985 principle of "Irrelevance of Remote Syntactic Structure"). According to this principle, no backward interaction between phonology and any preceding component is permitted. Only surface

structure can function as syntactic conditioning in phonology. Syntactic information that does not appear on the surface (such as my VP phantom rules from 3.3) cannot play a role in phonology.

I assume that the same holds for a cliticization component. A highly autonomous module of clisis is one where no backward interaction is tolerated. Cliticization will not, then, refer to syntactic information other than that available at surface structure. This makes clitic percolation rules of the sort suggested by Zwicky and Pullum (1983) highly suspect. Such percolation rules have the potential to interact with, for example, phantom rules, which do not license phrase structures, but serve to predict other PS rules. I comment further on clitic percolation rule below in 7.1.3.

The only syntactic information accessible to cliticization (and external sandhi — cf. Kaisse 1985) is the following:

(1) constituency
c-command relations
head of constituent
margin of constituent

These are all, in the monostratal GPSG framework I adopt here, statable with reference the surface structure (plus the rules that license those surface structures), the only level of representation.[2]

7.1.3. Toward A Restrictive Theory Of Cliticization.

A highly restricted cliticization component is one that displays uniformity among its various principles. For this reason it is important to exclude syntactic power from the cliticization operation known as liaison. Otherwise we will find that some cliticizations have purely phonological function, others have both syntactic and phonological function.

I have in mind here a potential distinction between two kinds of cliticizations: one that corresponds to Zwicky and Pullum's (1983) simple cliticization, and one that corresponds to their special cliticization. The first type refers to clitics that have the same distributions as their free variants. All that is required for these clitics is a readjustment of the syntactic structure so that the word becomes clitic to some host. I have labeled this particular readjustment *liaison*.

The second type refers to clitics that have distributions different from their free variants, or have no free form at all. If special clitics are not positioned through syntactic operations, as I argue in 4.2, then the sentential and phrasal location of the clitics has to be posited for the cliticization component.

A bound word like *-hAn* will then require the clisis module to locate it after the first constituent of the clause in which it occurs. Cliticization recapitulates the LP statement that locates certain members of adverb class [41] after the constituent marked [FIRST]. This approach misses a generalization here -- second position is the same for both syntax and cliticization -- and forces us to posit some syntactic power to clisis operations.

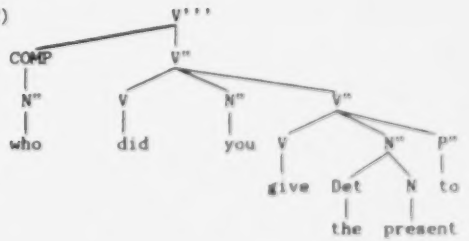
Only the first type of cliticization can be unified with external sandhi (and flattening). Such a tack is taken by Kaisse (1985), who groups special cliticization together with external sandhi, but keeps both separate from simple cliticization.

In this dissertation I have argued against the simple/special clitic division in favor of a bound word/phrasal affix approach. Under my view, only bound words are the result of a cliticization operation liaison. Phrasal affixes arise through a different source (and in a different component).

I offer a more restrictive theory of cliticization than Kaisse. My clisis module is, first of all, uniform. It is not split into two parts (cf. 7.2.2 below). There is only one type of operation that applies, phonological liaison. Syntactic operations are disallowed. As a phonological operation, cliticization can now be unified with external sandhi, which is likewise a set of phonological operations, sharing with liaison the fixed set of syntactic conditions and constraints in (1).

This component, which I term **readjustment**, satisfies all four of Zwicky's (1984c) and Sadock's (1983a) requirements for autonomy and limited interfacing. It is composed of liaison, external sandhi, and 'flattening'. The first two effect phonological linking between lexemes and phonological words; flattening reduces hierarchical syntactic structures to linear phonological structures. In English, for example, it linearizes a sentence like (2), so that the sentence will be grouped into a phrasing like (3), among other alternative phrasings (adopted from Zwicky 1982a).

(2)



(3) [whó did][you gíve][the présent to]

The readjustment module is uniform. All principles in it have to do with the match-up between syntax and phonology, especially phonological linking and grouping. There is no loss of uniformity as with Kaisse's splitting approach. It also has sharp boundaries. There seem not to be any principles intermediate between syntax and readjustment. Readjustment and syntax do not interact, except in limited acceptable backward interaction (conforming to Zwicky's PSCP). And finally, the readjustment component I offer here does not duplicate operations from syntax.

The syntactic operations that locate bound words are part of syntax proper, not achieved via liaison. If the readjustment module permits syntactic location of words (and/or morphemes), then syntax and readjustment will be coextensive in this regard. Such potential duplication of function points to a lost generalization, rather than (Sadock's) inherent overlap -- the two principles are wholly redundant and not simply partially redundant (see footnote 1).

The syntactic mechanisms needed for the positioning of bound words *-hAn*, *-µA*, *-kO*, and *-kin/-kAAn* are a proper subset of those mechanisms needed for the positioning of certain members of class [41] sentential adverbs. In effect these mechanisms are LP statements that refer to second position and to postverbal position. Duplication of function is avoided if we can use the same LP statements for the positioning of class [41] adverbs and for that needed for the positioning of the particle clitics.

7.2. The Interface Model.

The preceding discussion suffices to demonstrate that my readjustment component constitutes a highly restricted, autonomous module of grammar. In the following I couch my proposal in the broader schemes of Zwicky (1984c) and Kaisse (1985), with appropriate revisions.

7.2.1. Zwicky 1982b and 1984c.

Zwicky (1982b) sketches one proposal for the organization of modules in grammar. He presents the following scheme (repeated here from Zwicky 1984c).

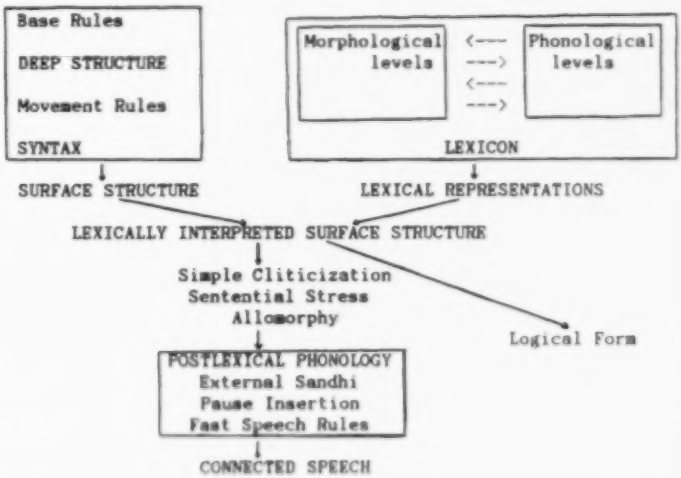
(2) relational (= cyclic) syntax
surface (= postcyclic) syntax
readjustment and cliticization
sentential prosody
free deletion
word formation
allomorphy
(nonautomatic) phonology (= morphonology)
surface filters
(automatic) phonology

Zwicky (1984c) later revises this format. First, he adopts a GPSG monostratal framework for syntax in place of a relational/surface (cyclic/postcyclic) division. Second, Zwicky is also willing to give up his earlier component of free deletion and to redistribute these principles to allomorphy and morphophonemics.

Zwicky (1984c) maintains stoutly that syntax and cliticization are to be kept separate, and that cliticization and readjustment constitute a common module, which follows syntax and precedes the other components.

7.2.2. Kaisse 1985.

Kaisse's (1985) takes Zwicky's 1982b model as point of departure. She incorporates a Government and Binding syntactic framework, as well as a Lexical Phonology framework for morphology and phonology:



Kaisse's accepts Zwicky and Pullum's (1983) simple/special clitic distinction, with which I have taken issue, so she is compelled to try to locate the two in separate components. Yet her conclusion notes so many similarites between the two that a componential division is put in doubt.

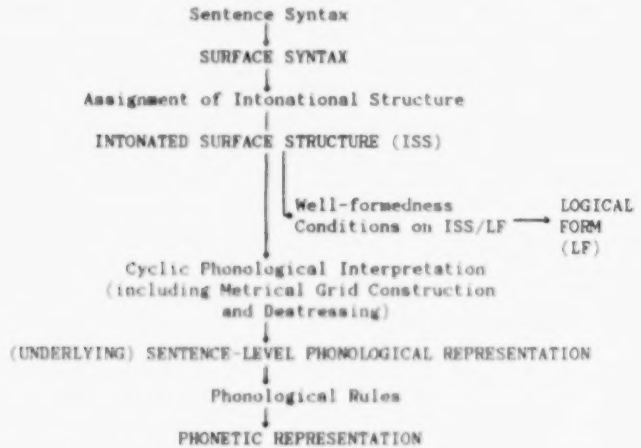
7.2.3. Selkirk 1984.

Selkirk's model of componential interaction is also amenable to my proposal. Her primary interest is the syntax-to-phonology mapping, a part of which is readjustment. Selkirk has two kinds of

cliticization, a syntactic cliticization and a rhythmic cliticization. Syntactic cliticization takes care of English *n't* and *to*-contraction, and is permitted to interact with other operations of syntax. Rhythmic cliticization covers the types of cliticization I have addressed in this dissertation. Her syntactic cliticization is dubious since Zwicky and Pullum (1983) have shown *n't* to be an affix. Likewise her *to*-contraction rule might also be affixation. And finally, some of the pronominal clitics might simply be prosodic leaners. In addition, her binary division of cliticization operations falls under the same criticisms as discussed in chapter 5: it fails to capture the right generalizations (with respect the the bound word/phrasal affix division) and further fails to provide a unified account of liaison.

Selkirk tends to lump together cliticization and prosodic operations. Her rule of AUX-contraction is one of a number of metrical grid construction rules (coupled with destressing). I have argued in chapter 4 that prosody and cliticization should not be grouped together in the same component (a similar point of view taken by Zwicky (1982a,b), for instance).

Selkirk's scheme is reproduced in the diagram below: (p. 34)

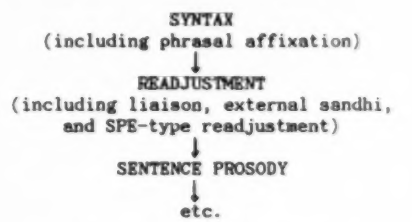


7.2.4. Revised Organization.

In this dissertation I have maintained that the syntactic and phonological aspects of cliticization are separate parameters. I have further argued that the syntactic positioning of clitics is a matter for syntax rather than cliticization. This approach reduces the notion

cliticization to mere liaison. Following Zwicky and Kaisse, we can now say that all of syntax precedes all of liaison.

Zwicky (1982b, 1984c) suggests that liaison (in his terms **simple cliticization**) is akin to readjustment; therefore he lumps the two together. I further suggest that external sandhi might belong in this module, as it exhibits uniformity with liaison operations (i.e. the conditions in (1)). I have also suggested that liaison takes precedence over sentence prosody. Hence I propose the following model of componential interfacing:



I will not take an unmotivated stand here as to the interfacing of these components with other components of the grammar, such as allomorphy. See instead Zwicky (1984c), Kaisse (1985), and Zwicky and Pullum (to appear).

The highly restricted model of interfacing and componential organization presented here is compatible with Klavans' arguments to the effect that syntactic and phonological parameters work independently of each other in cliticization.

7.3. Polysynthesis.

One potential extension of the framework that I have presented here is in the realm of polysynthesis. I have argued that clitics are either bound words or phrasal affixes. Bound words are lexemes marked with the lexical feature [± liaison]. I see no reason to limit the assignment of this feature to words belonging to nearly every word class, except lexical nouns and lexical verbs (Klavans 1982a). If we permit nouns and verbs to carry this feature, then we can connect bound words of the sort I have treated in this dissertation with noun incorporation and certain other sorts of polysynthesis. I am not claiming that all kinds of polysynthesis should be included here.

In particular, I have in mind an analysis of Greenlandic Eskimo by Sadock (1980), who distinguishes two types of object incorporation -- one where an inflected noun is incorporated (3), the other where a bare stem is incorporated (4).

(3) Palasip illua-nukarpoq.
priest/REL house/3SG-go to/INDIC/3SG
'He went to the priest's house.'

(4) Nuu-liarpoq
Godthaab-go to/INDIC/3SG
'He went to Godthaab.'

In (3), the possessed noun *illua* is incorporated into the verb; it is inflected with a possessive suffix. Its modifier *palasip* lies outside the incorporated noun + verb phonological word.

Sadock then suggests that

> The incorporation of inflected head nouns is, I suppose, closer to familiar sorts of cliticization than is the sort of incorporation that results in the omission of inflectional affixes. Yet it is strikingly different from typical cases of cliticization, as discussed by Zwicky 1977, in two major respects. First, there is almost never a free form of the affix that is phonologically (or historically) related. [Footnote deleted.] Second, the process results in a change in the grammatical class of the base form. (Sadock 1980:315)

If I am correct in associating this first kind of Greenlandic incorporation with my liaison operation, then we need not expect that there will be free forms for every word or morpheme nor that any change in word class will take place. The Finnish clitics *-hAn, -pA, -kO,* and *-kin/-kAAn* lack free forms entirely. Also the result of the attachment of the Finnish lexemes is not a change in word class, since liaison is a phonological operation, not a syntactic operation.

My suspicion is confirmed by Sadock's more recent work. Sadock (1983b) indicates a change in his approach to noun incorporation. He specifically refers to at least some of the more productive verb hosts for the incorporated nouns as clitics -- apparently clitics of the bound word variety rather than phrasal affixes.

I believe that a complex interaction of liaison and allomorphy is at work in Greenlandic Eskimo. I leave this topic open for pursuit at some later date. For now, however, I note that my framework, sketchy as it may be, seems sufficiently flexible for extension to such 'postcyclic' incorporation rules as those proposed by Sadock (though not for his 'precyclic', or lexical, incorporations). Since GPSG does not admit transformational rules -- hence there is no cyclic/postcyclic rule division -- Sadock's rules must belong to either syntax, or as he apparently now believes, cliticization.

7.4. Concluding Remarks.

I have, in course of this dissertation, argued for a strong theory of cliticization. In the case of the bound word, cliticization can be reduced to liaison, which is a readjustment operation that realigns syntactic structure in order to create phonological words. It lacks syntactic powers, but not syntactic conditions on its application. In the case of the phrasal affix, cliticization was shown to be a syntactic operation that locates affixes at the margins of constituents or beside heads.

A byproduct of my analyses is the addition of some proposals to GPSG concerning reference to margins of constituents, the notion of immediate adjacency, the VP phantom rules in Finnish, and the branching operation for phrasal affixes. I argue that reference to margins, as well as heads, is a necessary option for syntax (especially the immediate dominance part of grammar). The other part of GPSG, the LP format, proved too poor to handle some of the adverb positioning in Finnish, so I have proposed that the notion immediate adjacency be added to LP. A VP constituent is needed for some aspects of Finnish syntax, obtrusive for other aspects (such as free constituent order). Consequently I have made use of phantom VP rules, which do not directly license phrase structure trees, but instead serve to predict actual ID rules of the language. And finally, my discussion of the phrasal affix, although initially supported by existing analyses in GPSG, required additional branching in ID.

My proposals are, of course, based on the detailed description of only five morphemes in Finnish: *-hAn, -pA, -kO,* and *-kin/-kAAn* (coupled with somewhat less detailed descriptions of other data). I believe that the grammatical organization of Finnish is unlikely to differ in any fundamental way from that of any other language; so, my proposals entail a universalist approach to the phenomenon of cliticization. To this purpose, I have already suggested some possibilities for research in English and Greenlandic Eskimo.

Endnotes.

1. Sadock (1983a) argues that rules of distinct components sometimes have an inherent overlap in function -- without also being coextensive or complementary.

2. I say 'statable' here because these pieces of syntactic information are not necessarily stated in the phrase structure trees. The notion **head of constituent** is not actually mentioned as part of a phrase structure tree. In order to retrieve this information, one must take into account also the rules that license branching.

APPENDIX A

BOUND WORDS IN THE LEXICON

The bound words examined in the text are located syntactically before undergoing liaison. There is also a class of bound lexemes that appear only in compounds. Under my bound word analysis the bound compounding elements are lexemes that are marked [+LIAISON]. In all other respects these are regular lexemes in the lexicon and syntax.

In English one finds "cran(berry)" morphs (Aronoff 1976), which begin the word and lean rightward onto the host. In Finnish these are primarily leftward leaning words, often misconstrued as derivational suffixes. These bound stems appear to straddle the boundary between affix and lexeme, but it is important to examine this distinction because the morphemes in question attach outside inflectional affixes. Derivational affixes normally do not attach outside inflectional material. In (1) for example, *-läntä* 'rather, somewhat' attaches to *lihavan*, the genitive of *lihava* 'fat', stranding the inflectional suffix *-n* between the stem and the putative affix.

(1) lihavanläntä 'somewhat fat'

Several facts indicate that a bound word analysis is to be preferred over a derivational affix analysis. The solution I outline below avoids the malordering an affixal account encounters. The bound stems in question are *-lainen, -moinen, -mainen, -läntä,* and *-niekka.* They have two or more of the following word-like properties, which distinguish them, from regular affixes. All of them have the first two properties, and most of them have the other properties as well.

(2) Word-like properties of bound words:

A. attachment to a fully inflected form (NOM, GEN, PX) rather than the bare stem

B. violation of stress rules -- a cran-morph requires secondary stress on its initial syllable, regardless of which pattern is projected by the first stem

C. absence of vowel harmony with a preceding stem

D. permitting a particle clitic to intervene between the first stem and the second "cran" stem

E. attachment to whole phrases

F. conjunction reduction

Bound word *-lainen*.

The bound word *-lainen* 'type, sort, kind' is etymologically an agglutination of the word *#lajinen*, itself a derivation of *laji* 'type (plus derivational suffix *-nen*). My bound word account of *-lainen* recapitulates the history of the word. Bound word *-lainen* is not to be confused with derivational suffix *-lAinen*, which differs from the bound word homophone in most of the relevant properties.

Bound word *-lainen* does not attach to a bare stem as do regular derivational affixes, but to a fully inflected stem, in this case the genitive (3). Thus it requires that an inflectional affix (genitive *-n*) be sandwiched between it and a preceding stem. It also allows a possessive suffix to be stranded in the same position (4).

(3) hyvänlainen 'a good kind' (cf. hyvä 'good')
nykyisenlainen 'a modern kind' (cf. nykyinen 'modern')
eräänlainen 'a certain kind' (cf. eräs 'certin')
yhdenlainen 'homogeneous' (cf. yksi 'one')
tämänlainen[1] 'this kind' (cf. tämä 'this')

(4) hän on aivan entise-nsä lainen
s/he is quite former-3 kind
'She/He is quite like her/his former one' (cf. PX -nsa stranded between entinen 'former' and lainen)

-lainen further lacks vowel harmony with what precedes it. The examples in (3) show disharmony in vowel selection -- back vowel *a* cooccuring with front vowels *y* and *ä*. Thus one does not find *#hyvänläinen* etc.

This bound word permits a particle clitic to intervene between it and its host:

(5) toisenkinlainen -- toisenlainenkin 'another kind, too'
(cf. toinen '(an)other', particle clitic -kin 'also')

-lainen attaches to whole phrases:

(6) edellä mainitun lainen kirjailija
prec. mentioned kind writer
'a writer of the kind mentioned above

edellä sanotun lainen
prec. said kind
'of the sort just said'

Finally, *-lainen* exhibits a certain amount of syntactic freedom; it can participate in conjunction reduction (7).

(7) nykyisten ja entisten laisia autoja
modern and former type cars
'modern and former types of cars'

Thus *-lainen* has all of the six word-like properties listed above. This bound word is very productive and can combine with anything so long as that word or phrase is in the genitive. One apparent counterexample to this claim is *erilainen* 'different' (cf. *eri* 'separate'), but this occurs only because *eri* is an irregular adjective having no inflected forms, and thus *#erinlainen* would not be expected anyway. A few instances of bound word *-lainen* show lexicalization. Irregular *kahtalainen* 'of two kinds' and *yhtäläinen* 'identical, equal' have partitive hosts rather than genitive. *Kahtalainen* is a homonym of *kahdenlainen*. *Yhtäläinen*, however, is not the same as *yhdenlainen* 'homogeneous'. *Sekalainen* 'mixed, heterogeneous' has a nominative host, *seka-*.

As mentioned above, there is a homophonous derivational suffix *-lAinen* which lacks these properties. It has a different meaning, 'a person or thing from'; it attaches to nominatives (not genitives); it does not violates the stress rules, attach to phrases, undergo conjunction reduction, or permit a clitic to intervene between it and the stem; and it does show vowel harmony with the stem.

(8) helsinkiläinen 'someone from Helsinki' (*helsinginläinen)
suomalainen 'someone or something from Finland;Finnish'
(cf. nominative suomi, irregular stem suoma-)

Bound Word *-moinen*.

Bound word *-moinen* means 'such, like that, kind', and is similar to *-lainen*. Etymologically it comes from the stem *moko-* plus derivational suffix *-(i)nen*. *-moinen* attaches to inflected genitive stems and lacks vowel harmony (9), and permits intervening particle clitics (10). It also has a certain amount of phonological independence (11-12).

(9) (ei) minkäänmoinen '(not) any kind' (cf. minkään 'nothing')
*minkäänmöinen

(10) minkä-hän moinen 'what kind, I wonder' (cf. minkä irregular genitive of mikä 'what' and particle clitic -hän)

(11) moinen törkeys suorastaan kuvottaa

(12) en ole moista ennen nähnyt

-moinen is apparently not a true bound word, as the examples in (11) and (12) show. It is probably an unstressed word, a 'leaner' in Zwicky's (1982) terminology, and very likely not even an obligatory leaner -- just an optional leaner.

-moinen does not apparently permit conjunction reduction because of its restricted occurrence -- it is limited in its combinatory possibilities. A few irregularities are found in *-moinen* combinations. There are short stems of pronouns in *mimmoinen (minkämoinen) 'what kind', jommoinen (jonkamoinen) 'which kind', kummoinen 'what sort' (cf. stem ku-), tämmöinen (tämänmoinen) 'this kind'*. Vowel harmony occurs in this latter word, *tämmöinen*. There is assimilation of the final nasal *-n* to the initial *-m* of *-moinen* in all forms, including the pronominal stems above.

Bound word *-läntä*.

-läntä 'somewhat' attaches to genitives of adjectives (14). It violates word-internal vowel harmony, retaining its front vowels no matter what the preceding stem requires for harmony. *-läntä* is not a productive formant,[2] and does not attach to phrases, does not allow conjunction reduction, or permit a clitic to intervene between it and the stem (e.g. *#lyhyenkinläntä*).

(13) lyhyenläntä 'somewhat short, shortish' (lyhyt 'short')

Vähäläntä 'shortish, somewhat short' is irregular in having idiomatic meaning, cf. *vähän* 'few' (thus 'somewhat few' would be expected).

-läntä is claimed to be a deaffixation from denominal derivational suffixes *-lA* plus *-ntA* (Rintala 1972).

Bound Word *-niekka*.

Bound word *-niekka* shows possession. It is a borrowing from Russian *nik*. There is no vowel harmony with the host (14).

(14) häntäniekka kettu 'a fox having a tail'
tail-NIEKKA fox

-niekka attaches to nominative stems rather than genitive stems, compare nominative *häntä* 'tail' with (14). My Finnish informant tells me that *niekka* is not productive, and so does not combine with phrases, permit conjunction reduction, or allow a clitic to intervene between it and the stem.

Bound Word *-mAinen*.

-mAinen '-like' exhibits similar behavior to the bound words listed above. It attaches to nominatives (like *-niekka*, but in a productive manner, cf. (15)); it attaches to phrases and permits conjunction reduction (16); it agrees in vowel harmony with its host (17); and it permits/does not permit a clitic between it and its host.

(15) hammasmainen 'toothlike' (cf. hammas 'tooth')
brigittebardotmainen 'like Brigitte Bardot'

(16) puu- ja pensasmaisia kasveja 'tree and bush-like plants'
pöly- ja sumumaiset taistelukassut 'dust and fog-like combat gas'

(17) saksimainen 'scissor-like' -- *saksimäinen
käsimäinen 'hand-like' -- *käsimainen

Conclusion.

In summary I present a chart of the various word-like properties that this set of bound words can have: (the letters refer to the properties listed in (2)

	A	B	C	D	E	F
-lainen	+	+	+	+	+	+
-moinen	+	+	+	+	-	-
-läntä	+	+	+	-	-	?
-niekka	+	+	+	-	-	?
-mAinen	+	+	-	-	+	+

In the framework I present in chapter IV *-lainen, -läntä,* and *-niekka* are not, strictly speaking, bound words (phonologically attached lexemes), but are semi-clitics instead (prosodically subordinated words). *Moinen* is a mere optional leaner, able to stand alone sometimes. Only *-mAinen* shows the necessary phonological interaction with the host for inclusion in the class of bound words.

Rintala (1972:126) reports several other bound words of the type discussed above, including *-puoleinen, -sekainen, -sorttinen,*

-laatuinen, *-makuinen*, and -kokoinen. Karlsson (1983:193-4) offers examples of words that appear only in certain compound combinations: *irto-*, *-huiskea*, *jalko-*, *keski-*, *kiiras-*, *leivin-*, *lökä-*, *näsä-*, *pipar-*, *-pide*, *puos-*, *siite-*, *seko-*, *tierna-*, *toivio-*, *-tulkku*. These are for the most part unproductive stressless lexemes, sometimes borrowings from Swedish.

Karlsson's examples show that bound words and semi-clitics may lean rightward or leftward. Another rightward leaning compounding element is negative "prefix" *epä-*, which appears productively for example in *epäkohtelias 'impolite'*, *epämääräinen 'indefinite'*, *epä-suomalainen* 'un-Finnish', and the like. Although usually described as one of the only prefixes in Finnish, the following two facts are compatible with a semi-clitic analysis. *Epä-* does not agree in vowel harmony with its stem. It follows the regular stress patterns of a compound (and as the initial element it requires primary stress).

In conclusion, then, one finds in Finnish a number of bound words and semi-clitics in the lexicon. I assume that most of these have regular word properties, but are marked [+LIAISON] or [-STRESS], according to presence or absence of word-internal sandhi rules, respectively.

Endnotes.

1. Some irregularities may be noted here. The short form of *tämänlainen* is *tällainen* (even with vowel harmony in colloquial Finnish: *tälläinen*). Shortening, with subsequent assimilation of the *-n* of the genitive to *-l*, is typical of pronouns:

jonkalainen — jollainen
minkälainen — millainen

Assimilation occurs in spoken Finnish as a more general phenomenon than just the above three forms -- e.g. *kaikenlainen* 'all kinds' [kaikel:ainen].

Only irregular *yhtäläinen* 'identical, equal' and colloquial *tälläinen* exhibit vowel harmony with the host.

2. Rintala (1972) reports some dialectal usage of *läntä* that exhibits (a) productivity, (b) limited vowel harmony (e.g. *lapsenlanta* 'childish'), and (c) nominative hosts (e.g. *heikkoläntä* 'somewhat weak').

APPENDIX B

COMPARISON OF FINNISH AND LAPPISH MORPHOLOGY

I summarize Korhonen's (1981) main points of Lappish morphology and compare it to Finnish morphology.

a) In its nominal system Lappish exhibits three numbers (singular, dual, and plural) and eight cases. In comparison Finnish has only two numbers (singular and plural) but some thirteen productive cases:

	Lappish	Finnish
NUMBERS	SINGULAR	SINGULAR
	DUAL	---
	PLURAL	PLURAL
CASES	NOMINATIVE	NOMINATIVE
	GENITIVE	GENITIVE
	(ACCUSATIVE)	(ACCUSATIVE)
	ILLATIVE	ILLATIVE
		ALLATIVE
		INESSIVE
	LOCATIVE	ABESSIVE
		ELATIVE
		ABLATIVE
	COMITATIVE(SG)	COMITATIVE
	ABESSIVE	ABESSIVE
	ESSIVE (SG)	ESSIVE
	---	PARTITIVE
	---	TRANSLATIVE

Of particular interest is the lack of the partitive in Lappish, a prominent feature of Finnish morphosyntax. In Lappish only the "genitive-accusative" is used (i.e. in Northern Lappish the genitive), and there is no genitive/partitive division in the case marking of objects. Another source of divergence between the two is the local case system: the Lappish illative corresponds to the Finnish illative and allative; the Lappish locative (historically a mixture of the former inessive and elative cases) corresponds to the Finnish inessive, elative, adessive, and ablative (Korhonen 1981:221).

Finnish has a formally plural comitative, whereas Lappish has a singular comitative (with a clitic postposition in the plural, *-guim*). The essive in Lappish appears only in the singular, while Finnish has both a singular and a plural. The abessives in the two languages are rather different. Finnish has only the invariant morpheme *-tta(')/-ttä(')*; Northern Lappish has the clitic postposition *-tagu* (a desuffixation from caritive *#-ptu* + lative *#-g* + pleonastic lative *#(e)k* or *#(e)n*). Other Lappish languages and dialects exhibit various (clearly suffixal) forms of *#ptu*: Southern Lappish *-pth*, Western Lappish *-ptä*, Northern Lappish *-t'tu*. Also, there is no productive translative in Lappish as one finds in Finnish.

Both languages have an absolutive/possessive distinction. Lappish has a three-way enclitic system for its possessive declension (singular-dual-plural), whereas Finnish has only a two-way enclitic system (singular-plural).

In the adjective class, both languages have a positive/comparative/ superlative division, though the superlative is formally and etymologically different in the two languages: Finnish *-in:-imma-:-impa-* vs. Northern Lappish *-mus:-mussV-:-musâ-/-mos:-mčsV-:-bmusV-*. Furthermore Lappish displays a predicative/attributive distinction in its adjectives that does not occur in Finnish (where there is only one form, marked for case).

Finally, the adjective-noun concord systems are radically different in the two languages. Finnish has full concord; Lappish has the so-called "partial concord". In the Lappish system, most adjectives do not agree in number and case, but are instead in the special attributive (or actually base) form. Only certain word-forms permit "partial" agreement: the adjectives *buorre* 'good' and *bahâ* 'bad', the cardinal numbers, the ordinal *nub'be* 'second', and several pronouns. [Here there is no concord in the genitive, illative, or locative (in which the attribute is in the genitive singular); the attribute can be in the genitive plural in the genitive, illative, and comitative plural; in other cases the attribute agrees in number and case.]

b. The actual categories of verbal morphology are very similar in both languages. Both have four modes, two tenses, three persons, and numerous deverbal nominalized forms. Lappish and Finnish diverge, however, in their numbers -- as with the nominal morphology Lappish has three numbers (singular-dual-plural) and Finnish has only the two (singular-plural).

Both Finnish and Lappish have four verbal modes -- indicative (unmarked), potential, conditional, and imperative. The Northern

Lappish potential is formally different from that of Finnish (the Lappish potential being cognate with the Finnish conditional) and the Northern Lappish conditional is cognate with the Estonian and Livonian conditional *-ks* (though not with Finnish conditional *-isi-*).

The two languages have similar tense systems (present, past), though the morphological details differ somewhat. Likewise Finnish and Lappish share a three-person verbal system, though, again the details of form do not indicate much (recent) language contact.

In the nominalized forms, however, there are several divergences. The Lappish infinitive *-t* < **dĕk* is cognate with Finnish *-a/ä/ta/tä* < **-tak/täk*, both originally deverbal derivative **ta/tä* plus lative **-k*. The Northern Lappish "action" ending is *-m*, which is cognate with the Finnish third infinitive *-ma/mä*, but is usually translated into Finnish as the verbal noun in *-minen* or even derivational suffix *-o/ö*.

The first and second gerunds in Lappish are very distinct from Finnish. Northern Lappish has a first gerund *-ddiin/-dĕddiin* (< deverbal substantive formant **-nta/ntä* + plural *j* + locative **-na/nä*) where Finnish has the second infinitive inessive (infinitive *-te/de* + inessive *-ssa/ssä*). The second gerund in Northern Lappish is *-min* (<**mĕnĕ* < verbal noun **ma/mä* + essive-locative **na/nä*), which corresponds in meaning to Finnish first participle (*-va/vä*) or third infinitive in inessive or elative (*ma/mä* + *ssa/ssä* or *sta/stä*).

The Lappish verbal abessive is *-kaet'ta(i), -gnettĕ* (< deverbal **kka/kkä* + abessive **pta/ptä*). The Finnish abessive is *-matta/mättä* (= the third infinitive *ma/mä* + abessive). Both languages have cognate verbal adverbs -- NLp in *Ø*, F in *-en*.

The participles in the Lappish languages are rather different from Finnish. In the Northern Lappish present participle there is *Ø* (with a "contraction vowel"), *-g'gje* < **jĕ* < **ja/jä*. The Finnish correspondents are the present participle *-va/vä* (not etymologically the same as Lappish) and agentive *-ja/jä* (etymologically the same). The NLp perfect participle is *-m, -mas* < **-me*; the Finnish perfect participle is *-nut/nyt*.

In summation, then, all the inflectional categories and morphemes in Northern Lappish are due to genetic inheritance rather than linguistic interference from Finnish.

REFERENCES

Alvre, Paul. 1976. Vana kirjakeele küsusõnu. Keel ja Kirjandus 1976.6.

Alvre, Paul. 1981. *veelaks*-tüüp kirjakeeles ja murdeis. Keel ja Kirjandus 1981.1.

Anderson, Stephen R. 1984. Kwakwala Syntax and the Government-Binding Theory. In Syntax and Semantics 16:21-75. (The Syntax of Native American Languages.) Ed. by Eung-Do Cook and Donna B. Gerdts.

Ariste, Paul. 1973. Eesti rõhumäärsõna *ep*. Suomalais-Ugrilaisen Seuran Aikakauskirja 72:33-37.

Aronoff, Mark. 1976. Word Formation in Generative Grammar. The MIT Press, Cambridge, Massachusetts.

Baltin, Mark. 1982. A Landing Site Theory of Movement Rules. Linguistic Inquiry 13.1:1-38.

Beronka, Johan. 1922. Syntaksyiske iagttagelser fra de finske dialekter i Vadsjø og Porsanger. Videnskapsselskapets Skrifter II. Hist.-filos. Klasse. 1921. No. 10. Kristiania (Oslo), Norway.

Bissantz, Annette S. 1985. Syntactic Conditions on Two Types of English Cliticization in GPSG. OSU WPL 31:64-121.

Borer, Hagit. 1984. Parametric Syntax. Case Studies in Semitic and Romance languages. Studies in Generative Grammar 13. Foris Publications, Cinnaminson, New Jersey.

Breckenridge, Janet, and Auli Hakulinen. 1976. Cycle and after. Parassession on Diachronic Syntax, pp. 50-68. Chicago Linguistic Society.

Browne, W. 1974. On the Problem of Enclitic Placement in Serbo-Croatian. In Slavic Transformational Syntax, ed. by Richard D. Brecht and Catherine V. Chvany, pp. 36-52.

Carlson, Gregory N. 1983. Marking Constituents. In Linguistic Categories: Auxiliaries and Related Puzzles I:69-98. Ed. by F. Heny and B. Richards. Foris Publications, Dordrecht.

Collinder, Bjorn. 1949. The Lappish Dialect of Jukkasjärvi. Skrifter utgivna av K. Humanistiska Vetenskaps-Samfundet i Uppsala 37.3, Uppsala.

Comrie, Bernard. 1981. Morphology and Word Order Reconstruction: Problems and Prospects. In Historical Morphology, pp. 83-96, edited by Jacek Fisiak. Mouton.

Cowan, J. Ronayne. 1984. Evidence for the Aux Cliticization Hypothesis from Nuer. CLS 20:34-45.

Dixon, R.M.W. 1980. The Languages of Australia. Cambridge Language Surveys. Cambridge University Press.

Dressler, Wolfgang U. 1985. Morphonology: The Dynamics Of Derivation. Karoma Publishers, Ann Arbor.

Fokos-Fuchs, D. R. 1962. Rolle des Syntax in der Frage nach Sprachverwandtschaft. Ural-Altaische Bibliothek XI. Otto Harrassowitz, Wiesbaden.

Fuchs, D. R. 1937. Übereinstimmungen der syntax der finnisch-ugrischen Sprachen. FUF XXIV:292-322.

Fuchs, D. R. 1949. Der Komparativ und Superlativ in den finnisch-ugrischen Sprachen. FUF XX:147-230.

Fudge, E. C. 1969. Syllables. Journal of Linguistics, pp. 253-286.

Fujisaki, Hiroya, and Ilse Lehiste. 1983. Some Temporal Characteristics of Declarative Sentences in Estonian. Preprints of Papers for the Working Group on Intonation, The XIIIth International Congress of Linguists, Tokyo, 1982.

Gazdar, Gerald. 1980. Phrase Structure Grammar. In G. K. Pullum and P. Jakobson (eds.) On the Nature of Syntactic Representation.

Gazdar, Gerald. 1984. The Head Feature Convention as a Default Mechanism. Unpublished paper presented at the Annual Meeting of the LSA, Baltimore. This paper will be incorporated into Gazdar, Klein, Pullum, and Sag (in press).

Gazdar, Gerald, Geoffrey K. Pullum, and Ivan Sag. 1982. Auxiliaries and related phenomena in a restrictive theory of grammar. Language 58.3:591-638.

Gazdar, Gerald, Geoffrey K. Pullum, Ivan Sag, and Ewan Klein. 1985. Generalized Phrase Structure Grammar. Harvard University Press.

Genetz, Arvid. 1892. Suomen partikkelimuodot. Suomalaisen Kirjallisuuden Seura, Helsinki.

Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's field trip. CLS 7:394-415.

Groos, Anneke. 1978. Towards an Inflectional Theory of Clitics. Manuscript, University of Amsterdam.

Hakulinen, Auli. 1975. Suomen *sitä*: pragmatiikan heijastuma syntaksissa. [English summary: Finnish *sitä* -- an Instance of the Interplay between Syntax and Pragmatics.] Sananjalka 17:25-41.

Hakulinen, Auli. 1976. Liitepartikkelin *-han/-hän* syntaksia ja pragmatiikkaa. Reports on text linguistics: Suomen kielen generatiivista lauseoppia 2. (Meddelanden från stiftelsens för Åbo Akademi Forskningsinstitut, no. 7:3-60.)

Hakulinen, Auli. 1984a. On Cohesive Devices in Finnish. In Text Connexity, Text Coherence, ed. by E. Sözer. (Papers in Text Linguistics.) Busche Verlag, Hamburg.

Hakulinen, Auli. 1984b. Text Linguistics in Finland. Manuscript, University of Helsinki.

Hakulinen, Auli, and Fred Karlsson. 1979. Nykysuomen lauseoppia. Suomalaisen Kirjallisuuden Seuran toimituksia 350, Helsinki.

Hakulinen, Auli, Fred Karlsson, and Maria Vilkuna. 1980. Suomen tekstilauseiden piirteitä: Kvantitatiivinen tutkimus. Department of General Linguistics, University of Helsinki, Publications No. 6.

Hakulinen, Lauri. 1979. Suomen kielen rakenne ja kehitys. Werner Söderström Osakeyhtiö, Porvoo. (Originally published in 1957-59.)

Hale, Kenneth. 1973. Person marking in Warlbiri. In A Festschrift for Morris Halle, edited by Stephen R. Anderson and Paul Kiparsky, pp. 308-44. Holt, Rinehart and Winston, New York.

Hale, Kenneth. 1981. On the Position of Warlbiri in a Typology of the Base. Distributed by the Indiana University Linguistics Club, Bloomington.

Hock, Hans. 1982. Aux cliticization as a motivation for word order change. Studies in the Linguistic Sciences 12.1:91-101.

Itkonen, Erkki. 1960. Lappische Chrestomathie mit grammatikalischem Abriss und Wörterverzeichnis. Hilfsmittel für das Studium der finnisch-ugrischen Sprachen VII. Suomalais-Ugrilainen Seura, Helsinki.

Itkonen, Terho. 1981. Lapsi suomen taivutussuffiksien oppijana. (Review of Jorma Toivainen 1980) Virittäjä 85.2:152-156.

Jaeggli, Osvaldo. 1982. Topics in Romance Syntax. Studies in Generative Grammar 12. Foris Publications, Dordrecht, Holland/Cinnaminson, New Jersey.

Janda, Richard. 1981. A Case of Liberation from Morphology into Syntax: The Fate of the English Genitive-Marker *-(e)s*. Syntactic Change. Natural Language Studies 25:59-114.

Janda, Richard. 1984. Why Morphological Metathesis-Rules Are Rare: On the Possibility of Historical Explanation in Linguistics. BLS 10.

Janson, Tore. 1976. Placement of Enclitics in Latin and the Relation between Syntax and Phonology. Papers from the Third Scandinavian Conference of Linguistics, pp. 237-44. Edited by Fred Karlsson.

Jeffers, Robert J., and Arnold M. Zwicky. 1980. The Evolution of Clitics. Papers from the Fourth International Conference on Historical Linguistics, pp. 55-64, ed. by E. Traugott et al. John Benjamins. Also in OSU WPL 24:55-64.

Joseph, Brian D. 1983. The Synchrony and Diachrony of the Balkan Infinitive. Cambridge Studies in Linguistics, Supplementary volume. Cambridge University Press.

Kaisse, Ellen M. 1981. Luiseño particles and the universal behavior of clitics. Linguistic Inquiry 11:422-34.

Kaisse, Ellen M. 1982. Sentential Clitics and Wackernagel's Law. Proceedings of the First West Coast Conference on Formal Linguistics, edited by Daniel P. Flickinger, Marlys Macken, and Nancy Wiegard. Pp. 1-14.

Kaisse, Ellen M. 1983. The English auxiliaries as sentential clitics. In Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax, pp. 96-102. Ed. by John F. Richardson, Mitchell Marks, and Amy Chukerman. Chicago Linguistic Society.

Kaisse, Ellen M. 1985. Connected Speech: The Interaction of Syntax and Phonology. Academic Press, Orlando.

Karlsson, Fred. 1983. Suomen kielen äänne- ja muotorakenne. Werner Söderström Osakeyhtiö, Porvoo.

Karlsson, Fred. 1984a. Tagging and Parsing Finnish. Manuscript, University of Helsinki.

Karlsson, Fred. 1984b. Paradigms and Word Forms. To appear in Papers from the Workshop on Natural Morphology, ed. by Roman Laskowski.

Karlsson, Fred, and Benny Brodda. 1981. An Experiment with Automatic Morphological Analysis of Finnish. Department of General Linguistics, University of Helsinki, Publications No. 7.

Karttunen, Frances. 1975a. The Syntax and Pragmatics of the Finnish Clitic *-han*. Texas Linguistic Forum 1:41-50.

Karttunen, Frances. 1975b. More Finnish Clitics: Syntax and Pragmatics. Reproduced by the Indiana University Linguistics Club.

Karttunen, Frances. 1975c. Functional Constraints in Finnish Syntax. In Papers from the Parasession on Functionalism, ed. by Robert E. Grossman, L. James San, and Timothy J. Vance. Chicago Linguistic Society.

Klavans, Judith. 1979. On Clitics as Words. The Elements: A Parasession on Linguistic Units and Levels, Chicago Linguistic Society.

Klavans, Judith. 1982a. Some Problems in a Theory of Clitics. Distributed by the Indiana University Linguistics Club. 1980 Dissertation, University College London.

Klavans, Judith. 1982b. Configuration in a Non-configurational Language. WCCFL 1:292-306.

Klavans, Judith. 1983. The morphology of cliticization. In Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax, pp. 103-121. Ed. by John F. Richardson, Mitchell Marks, and Amy Chukerman. Chicago Linguistic Society.

Klavans, Judith. 1985. The Independence of Syntax and Phonology in Cliticization. Language 61.1:95-120.

Korhonen, Mikko. 1974. Oliko suomalais-ugrilainen kantakieli agglutinoiva? Eli mitä kielihistoriallisista rekonstruktioista voidaan lukea ja mitä ei. Virittäjä 1974:243-254.

Korhonen, Mikko. 1981. Johdatus lapin kielen historiaan. Suomalaisen Kirjallisuuden Seuran toimituksia 370, Helsinki.

Koski, August A., and Ilona Mihalyfy. 1964. Hungarian Basic Course II. Foreign Service Institute Basic Course Series. Department of State, Washington, D.C.

Kuznetsov, P. S. 1960. Russkaya Dialektologiya. Moskva.

Larjavaara, Matti. 1979. Lyydin *se*-pronomini ja -partikkeli. Sivulaudatur-työ, University of Helsinki.

Lehiste, Ilse. 1970. Suprasegmentals. MIT Press.

Lehiste, Ilse. 1982a. Phonetic Manifestation of Syntactic Structure in English. The Tokyo Journal of Medical Sciences, vol. 89.1/2:1-22.

Lehiste, Ilse. 1982b. Some Phonetic Characteristics of Discourse. Studia Linguistica 36.2:117-130.

Lehiste, Ilse. 1983a. The Role of Prosody in the Internal Structuring of a Sentence. Proceedings of the XIIIth International Congress of Linguists, Tokyo.

Lehiste, Ilse. 1983b. Signalling of Syntactic Structure in Whispered Speech. Folia Linguistica XVIII.1/2:239-245.

Leino, Pentti. 1982. Kieli, runo, ja mitta. Suomalaisen Kirjallisuuden Seuran toimituksia 376. Suomalaisen Kirjallisuuden Seura, Pieksämäki.

Matthews, Peter. 1972. Inflectional Morphology: A Theoretical Study Based On Aspects Of Latin Verb Conjugation. Cambridge University Press, Cambridge.

Matthews, Peter. 1974. Morphology: An Introduction to the Theory of Word-Structure. Cambridge University Press Cambridge.

Nevis, Joel A. 1981. Object Case Marking in Finnish. Unpublished M.A. thesis, OSU.

Nevis, Joel A. 1982. Suffix Versus Clitic: Four Estonian Case Endings. CLS 18:400-406.

Nevis, Joel A. 1984a. Possessive Suffix or Anaphoric Clitic: Five Finnish Morphemes. OSU WPL 29:174-207.

Nevis, Joel A. 1984b. The Clitic as Deviation from the Prototypical Word. Paper presented at the Annual Meeting of the Linguistic Society of America, Baltimore.

Nevis, Joel A. 1984c. On Finnish Disyllables, Cliticization, and Typology. Paper at the Third Uralic Phonology Conference (Eisenstadt, Austria). Paper to appear in conference proceedings.

Nevis, Joel A. 1984d. A Non-Endoclitic in Estonian. Lingua 64.2/3:209-224.

Nevis, Joel A. 1985a. Current Topics in Estonian Clisis. Presentation at the Third Annual Meeting of the Finno-Ugric Studies Association of Canada, Montreal.

Nevis, Joel A. 1985b. Decliticization in Old Estonian. To appear in OSU WPL 34 (1986).

Nevis, Joel A. 1985c. Language External Evidence for Clitics as Words: Lappish Particle Clitics. CLS 21:289-305.

Nielsen, Konrad. 1979. Lapp Dictionary/Lappisk (Samisk) Ordbok, Second edition. Volume I: A-F. Universitetsforlaget, Oslo.

Niemi, Jussi. 1984. Word Level Stress and Prominence in Finnish and English: Acoustic Experiments on Production and Perception. University of Joensuu Publications in the Humanities, No. 1.

Nykysuomen Sanakirja (NSSK = Dictionary of Modern Finnish). Werner Söderström Osakeyhtiö, Porvoo.

Ojansuu, Heikki. 1922. Itämerensuomalaisten kielten pronominioppia. Turun Suom. Yliopiston julkaisuja, B.1.3. Turku.

Östman, Jan-Ola. 1977. On English *also*, Finnish *-kin*, and related matters: A contrastive approach. In Contrastive Linguistics and Error Analysis, ed. by R. Palmberg and H. Rinbgom. Publications of the Research Institute of the Åbo Akademi Forskningsinstitut, no. 19.

Penttilä, Aarni. 1957. Suomen kielioppi. Werner Söderström Osakeyhtiö, Porvoo.

Pulgram, Ernst. 1970. Syllable, Word, Nexus, Cursus. [Janua Linguarum 81.] Mouton, The Hague/Paris.

Ravila, Paavo. 1941. Über die Verwendung der Numeruszeichen in den uralischen Sprachen. FUF XXVII.1-3:1-136

Rintala, Päivi. 1972. Suomen *länttä*-loppuiset adjektiivit. Suomalaisen Kirjallisuuden Seuran toimituksia 306.

Sadeniemi, Matti. 1949. Metriikkamme perusteet (ja sovellutusta moderneihin ja antiikin mittoihin). Suomalaisen Kirjallisuuden Seuran Toimituksia 236. Otava, Helsinki.

Sadock, Jerrold M. 1980. Noun incorporation in Greenlandic. Language 56.2:300-319.

Sadock, Jerrold M. 1983a. Linguistics as Anatomy. Presentation to the OSU Linguistics Department.

Sadock, Jerrold M. 1983b. The necessary overlapping of grammatical components. In Papers from the Parasession on the Interplay of Phonology, Morphology, and Syntax, pp. 198-221. Ed. by John F. Richardson, Mitchell Marks, and Amy Chukerman. Chicago Linguistic Society.

Sammallahti, Pekka. 1977. Norjansaamen Itä-Enontekiön murteen äänneoppi. Suomalaisen Kirjallisuuden Seuran toimituksia 160, Helsinki.

Saxon, Leslie. 1985. Agreement in Dogrib: Inflection or Cliticization? Paper presented at the Symposium on Canadian Native Languages in Theoretical Perspective, SUNY at Buffalo.

Schachter, Paul. 1974. Constraints on Clitic Order in Tagalog. UCLA Papers in Syntax 5.

Selkirk, Elisabeth O. 1972. The Phrase Phonology of English and French. PhD dissertation, MIT.

Selkirk, Elisabeth O. 1980. Prosodic Domains in Phonology: Sanskrit Revisited. In Juncture, ed. by Mark Aronoff and Mary-Louise Kean. Anma Libri, Saratoga, Caflifornia.

Selkirk, Elisabeth O. 1984. Phonology and Syntax: The Relation Between Sound and Structure. MIT Press, Cambridge, Massachusetts.

Setälä, E. N. 1930. Suomen kielen lauseoppi. Otava, Helsinki.

Siro, Paavo. 1977. Sijakielioppi. Oy Gaudeamus Ab, Helsinki.

Smith, Ian, and Steve Johnson. 1984. Clitics as a derived category. Paper presented at the Annual Winter Meeting of the Linguistic Society of America, Baltimore.

Steele, Susan. 1976. On the Count of One. In Studies Presented to Joseph Greenberg, ed. by Alphonse Juilland. Anma Libri and Company, pp. 591-613.

Steele, Susan. 1977. Clisis and Diachrony. In Mechanisms of Syntactic Change, pp. 539-79. Ed. by Charles N. Li. Blackwell, Oxford.

Steele, Susan. 1979. Uto-Aztecan: An Assessment for Historical and Comparative Linguistics. In The Languages of Native America: Historical and Comparative Assessment, pp. 444-544. Ed. by Lyle Campbell and Marianna Mithun. University of Texas Press, Austin.

Steele, Susan, et al. 1981. An Encyclopedia of AUX: A Study in Cross-Linguistic Equivalence. MIT Press.

Stump, Gregory. 1980. An Inflectional Approach to French Clitics. OSU WPL 26:1-54.

Tegey, Habibulah. 1975. A Study of Pashto Clitics and Implications for Linguistic Theory. Studies in the Linguistic Sciences 5.1:154-196.

Thomason, Sarah. 1978. Are There Linguistic Prerequisites for Contact-induced Language Change? Paper read at the University of Wisconsin-Milwaukee Symposium on Language Change.

Thomason, Sarah. 1980. Morphological Instability, With and Without Language Contact. In Historical Morphology, pp. 359-72. Ed. by Jacek Fisiak. [Trends in Linguistics: Studies and Monographs 17.] Mouton Publishers, The Hague.

Toivainen, Jarmo. 1980. Inflectional Affixes Used by Finnish-Speaking Children Aged 1-3 Years. Suomalaisen Kirjallisuuden Seuran toimituksia 359, Helsinki.

Toivainen, Jarmo. 1984. Kysymysmorfeemien ilmaantuminen suomalaislasten kielessä/The Acquisition of the Interrogative Morphemes of Finnish. Publications of the Department of Finnish and General Linguistics of the University of Turku 19.

Välimaa-Blum, Riitta. 1985. The Discourse Function of the Finnish Clitic *-han*: Another Look. Paper to be read at the Sixth International Conference of Nordic and General Linguistics, August 1986, Helsinki, Finland.

Vesikansa, Juoko. 1977. Johdokset. Nykysuomen oppaita 2. Werner Söderström Osakeyhtiö, Porvoo.

Vilppula, Matti. 1984. *Kurkikin jo lähti*. (On the Meaning of the Clitic *-kin*). Virittäjä 88.1.

Voegelin, C. F., and Voegelin, F. M. 1977. Classification and Index of the World's Languages. Elsevier, New York.

Vuoriniemi, Jorma. 1973. Ollin kieli nykysuomen kuvastimena. Suomalaisen Kirjallisuuden Seuran toimituksia 312.

Wackernagel, Jakob. 1892. Über ein Gesetz der indigermanischen Wortstellung. IF I(1892):333-436.

Zwicky, Arnold M. 1969. Phonological Constraints in Syntactic Description. Papers in Linguistics 1.3:411-63.

Zwicky, Arnold M. 1970. Auxiliary Reduction in English. Linguistic Inquiry 1.3:323-36.

Zwicky, Arnold M. 1977. On Clitics. Distributed by the Indiana University Linguistics Club, Bloomington.

Zwicky, Arnold M. 1982a. Stranded *to* and phonological phrasing in English. Linguistics 20.3-57.

Zwicky, Arnold M. 1982b. An Expanded View of Morphology in the Syntax-Phonology Interface. Preprints for the Plenary Session Papers, the XIIITH International Congress of Linguists. Tokyo, Japan. pp. 168-178.

Zwicky, Arnold M. 1984a. Clitics and Particles. OSU WPL 29:105-116. Also in Language 61.2:283-305.

Zwicky, Arnold M. 1984b. Welsh Soft Mutation and the Case of Object NPs. CLS 20:387-402.

Zwicky, Arnold M. 1984c. Autonomous Components and Limited Interfacing: Phonology-Free Syntax, the Hallean Syllogism, and Their Kin. CLS 20:365-386.

Zwicky, Arnold M. 1985a. How to Describe Inflection. BLS 11.372-86.

Zwicky, Arnold M. 1985b. Margins and Heads in GPSG. Abstract for a paper.

Zwicky, Arnold M. 1985c. Free Word Order in GPSG. OSU WPL 32 (1986), pp. 125-132.

Zwicky, Arnold M. 1985d. Heads. Journal of Linguistics 21:1-29. Revised version from OSU WPL 29:59069.

Zwicky, Arnold M. and Geoffrey Pullum. 1983. Cliticization vs. Inflection: English n't. Language 59:502-513.

Zwicky, Arnold, and Geoffrey Pullum. 1984. The Syntax-Phonology Boundary and Current Syntactic Theories. OSU WPL 29:105-116.

Zwicky, Arnold, and Geoffrey Pullum. To Appear. The Syntax-Phonology Interface. Academic Press.

Additional Reference:

Derbyshire, Desmond C. 1979. Hixkaryana. Lingua Descriptive Series I. North Holland Publishing Co., Amsterdam.

ABBREVIATIONS:

BLS = The Proceedings of the Annual Meetings of the Berkeley Linguistic Society

CLS = The Proceedings of the Annual Meetings of the Chicago Linguistic Society

FUF = Finnish-Ugrische Forschungen

IF = Indogermanische Forschungen

OSU WPL = The Ohio State University Working Papers in Linguistics

WCCFL = The Proceedings of the West Coast Conference on Formal Linguistics